Multivariable Calculus

Thomas Kong

July 2, 2022

Contents

Chapter 1

Preliminaries

1.1 Vector Spaces

Our end goal is to do *multivariable* calculus. That is, we want to differentiate and integrate (real) vector-valued functions of several variables. To this end, we'll need to rigourously understand what a vector is so that we can comfortably manipulate them (beyond the highschool 'definition' of 'a vector is a quantity with both a magnitude and a direction'). Thus, we'll begin with a lightning review of basic concepts from linear algebra, beginning with the *actual* definition of a vector. A vector is simply an element of a structure called a *vector space*, which we define below:

Definition 1.1.1. A **vector space** over a field $\mathbb{K}$ is a set V, together with the operations of 'vector addition' $f : V \times V \to V$ and 'scalar multiplication' $g : \mathbb{K} \times V \to V$, typically denoted by $f(x, y) = x + y$ and $g(\alpha, x) = \alpha x$, which satisfy the following axioms:

(a) V is an abelian group with respect to the binary operation of vector addition.

(b) Associativity of scalar multiplication: For every $v \in V$ and $a, b \in \mathbb{K}$, $(ab)v = a(bv)$.

(c) Distributivity of scalar multiplication: For every $v \in V$ and $a, b \in \mathbb{K}$, $(a + b)v = av + bv$.

(d) Distributivity of scalar multiplication over vector addition: For every $v, w \in V$ and $a \in \mathbb{K}$, $a(v + w) = av + aw$.

(e) Multiplicative identity: For every $v \in V$, $1v = v$.

We will only consider vector spaces over the field of real numbers; from now on we will let the arbitrary field $\mathbb{K}$ be $\mathbb{R}$, and herein when we say 'vector space' we are implicitly referring to a *real* vector space. The prototypical example of a vector space is the very first vector space that we all worked with before we even heard the term 'vector space': namely, n-tuples of real numbers, $\mathbb{R}^n$. Indeed, this will basically be the only vector space we will care about, as we will later see.

We'll now fly through some standard linear algebra concepts. Let V be a vector space. A **subspace** is a subset $S \subset V$ such that S is itself a vector space under the same vector addition and scalar multiplication operations, restricted to S. Fortunately, to check that a subset of a vector space is a subspace, one does not actually have to check every single one of the vector space axioms. A necessary and sufficient condition for this is simply for the subset to be closed under vector addition and scalar multiplication, i.e. for all $x, y \in S$ and $a \in \mathbb{R}$, $x + y \in S$ and $ax \in S$.

A mapping $T : V \to W$ between vector spaces that preserves the vector space structure is called a **linear map** or a **vector space homomorphism**, i.e. for all $x, y \in V$ we have $T(x + y) = T(x) + T(y)$ and for all $a \in \mathbb{R}$ we have $T(ax) = aT(x)$. A bijective homomorphism is called an **isomorphism**. Two vector spaces are *isomorphic* is there exists an isomorphism between them. Isomorphic vector spaces are essentially 'the same' for all intents and purposes; vector spaces that are isomorphic share all the same algebraic properties.

Let $T : V \to W$ be a vector space homomorphism. The set $\ker T = \{x \in V \mid T(x) = 0\}$ is called the **kernel** or **null space** of T, and the set $\mathcal{R}(T) = \{T(x) \mid x \in V\}$ is called the **range** of T. It can be easily shown that $\ker T$ is a subspace of V and $\mathcal{R}(T)$ is a subspace of W. Another useful fact worth noting is that a linear map T is injective if and only if its kernel is trivial, i.e. $\ker T = \{0\}$ (see exercises). Hence, T is an isomorphism if and only if $\ker T = \{0\}$ and $\mathcal{R}(T) = W$.

A particularly important way that we use to classify vector spaces is the notion of dimension, which intuitively speaking is the 'number of degrees of freedom' it possesses. We will proceed to formalise this below. To find out how many degrees of freedom a vector space has, we essentially need to find the minimum number of fixed vectors required to write any arbitrary vector from the space as some weighted sum of these fixed vectors. We call such weighted sums of vectors **linear combinations**; a linear combination of the vectors $v_1, \ldots, v_n$ is a sum $\sum_{i=1}^{n} c_i v_i$ for some constants $c_i \in \mathbb{R}$. The **span** of a subset A of some vector space V is the set of all possible linear combinations of finite subsets of vectors of A. Hence, in order to describe the entire vector space in terms of sums of vectors from one of its subsets A, we require that $V = \text{span} A$.

However, even if we are able to identify a subset $A \subset V$ such that $V = \text{span } A$, it is possible that we are able to find a smaller subset that does the trick, which suggests that the original subset may contain redundant information. The way we describe this redundancy is through the concept of linear dependence:

Definition 1.1.2. Let V be a vector space. A subset $A \subset V$ is **linearly independent** if for every finite subset $\{v_1, \ldots, v_n\}$ of A, we have that $a_1 v_1 + \cdots + a_n v_n = 0$ for scalars $a_1, \ldots, a_n \in \mathbb{R}$ implies that $a_1 = \cdots = a_n = 0$.

If A is not linearly independent, then it is **linearly dependent**.

At last, we arrive at our desired criteria for a subset of a vector space to summarise all of the information of V in the most minimalistic way possible:

Definition 1.1.3. Let V be a vector space. A **basis** for V is a linearly independent subset $A \subset V$ such that $V = \text{span } A$.

'Fun' remark: The definition we provide above is that of a *Hamel* basis. There are other types of bases. However, this distinction is not relevant for us in the finite dimensional setting, which we will be exclusively working in. The problem arises in infinite dimensional vector spaces. While the Axiom of Choice guarantees the existence of a Hamel basis for an arbitrary vector space, the basis may be uncountable. Seeing that we use bases to describe a vector space in terms of a reduced amount of vectors, having an uncountable basis means that there is still a 'large amount of information' needed to describe the vector space. This will not be a problem for us, however.

So, we would like to say that the number of elements in a basis quantifies the number of degrees of freedom that the vector space possesses. However, one concern that arises is whether in our definition above that there is a possibility that there are bases with different numbers of elements. Fortunately, the answer is no, as we will show shortly. Before we proceed with the proof of this particular fact, we will first consider an intermediate result that will greatly simplify the later proof.

Lemma 1.1.4. *(Replacement Lemma) Let V be a vector space. Suppose A is a linearly independent subset of V consisting of m elements and suppose B is a subset of V consisting of n elements such that $\text{span } B = V$. Then*

(a) There exists a proper subset $B' \subset B$ consisting of $n - m$ elements such that $\text{span}(A \cup B') = V$.

(b) $m \leq n$.

Proof. Let $A = \{x_1, \ldots, x_m\}$ and $B = \{v_1, \ldots, v_n\}$. We will prove that for each $k \in \{0, \ldots, m\}$, we have that $k \leq n$ and that the set $A_k \cup B'_k$ spans V where

$$A_k = \{x_1, \ldots, x_k\}$$

and

$$B'_k = B \backslash \{v_1, \ldots, v_k\}$$

after possibly some relabelling of the vectors v_i.

First, consider the case $k = 0$. So $A_0 = \emptyset$ and $B_0' = B\backslash\emptyset = B$. We thus have that $\text{span}(A_0 \cup B_0') = \text{span } B = V$ and $n \geq 0 = k$. This proves the claim for $k = 0$.

Now suppose that the statement holds for some $k < m$. We will show that the statement must also hold for the $k+1$ case. By hypothesis, we have that $k \leq n$ and

$$\text{span}(A_k \cup B_k') = \text{span}\{x_1, \ldots, x_k, v_{k+1}, \ldots, v_n\} = V.$$

Assume for a contradiction that $k+1 > n$. Combined with the fact that $k \leq n$, we must have that $k = n$, so that $B_k' = \emptyset$. Hence,

$$\text{span}(A_k \cup B_k') = \text{span } A_k = V.$$

This implies that the vector $x_{k+1} \in V$ can be expressed as a nonzero linear combination of the vectors in A_k. This would, however, contradict the linear independence of the set A_{k+1}. We must therefore have that $k+1 \leq n$.

We will now explicitly express x_{k+1} as a linear combination of the vectors in the set $A_k \cup B_k'$, noting that now $B_k' \neq \emptyset$

$$x_{k+1} = a_1 x_1 + \ldots a_k x_k + a_{k+1} v_{k+1} + \cdots + a_n v_n.$$

Since $x_{k+1} \neq 0$, then not all of the coefficients a_i are zero. In fact, we must have that $a_i \neq 0$ for some $i \geq k+1$ (if this were not the case, it would follow that x_{k+1} is a linear combination of the vectors $x_1, \ldots, x_k$, which would contradict the linear independence of the set A_{k+1}). Without loss of generality, we can assume that $a_{k+1} \neq 0$. Hence, we can write

$$v_{k+1} = \frac{1}{a_{k+1}}(x_{k+1} - a_1 x_1 - \cdots - a_k x_k - a_{k+2} v_{k+2} - \ldots a_n v_n).$$

We'll now demonstrate that the set $A_{k+1} \cup B_{k+1}'$ spans V. Let $y \in V$. Writing y in terms of the spanning set $A_k \cup B_k' = \{x_1, \ldots, x_k, v_{k+1}, \ldots v_n\}$ yields

$$y = c_1 x_1 + \cdots + c_k x_k + c_{k+1} v_{k+1} + \cdots + c_n v_n$$

for some constants $c_1, \ldots, c_n \in \mathbb{R}$. Substituting in our expression for v_{k+1} gives

$$y = \left(c_1 - \frac{a_1}{a_{k+1}}\right)x_1 + \cdots + \left(c_k - \frac{a_k}{a_{k+1}}\right)x_k + \frac{c_{k+1}}{a_{k+1}}x_{k+1} + \left(c_{k+2} - \frac{a_{k+2}}{a_{k+1}}\right)v_{k+2} + \cdots + \left(c_n - \frac{a_n}{a_{k+1}}\right)v_n.$$

Hence $y \in \text{span}\{x_1, \ldots, x_{k+1}, v_{k+2}, \ldots, v_n\} = \text{span}(A_{k+1} \cup B_{k+1}')$. Since the choice of $y \in V$ was arbitrary, it follows that $V = \text{span}(A_{k+1} \cup B_{k+1}')$. This completes the proof of the claim for the $k+1$ case.

Thus, by induction on k, the claim is proven for all $k \in \{0, \ldots m\}$. The case $k = m$ is exactly the statement of the lemma, so we are done. $\square$

The Replacement Lemma will serve as a nice stepping stone towards an easy proof of the following crucial fact about bases:

Theorem 1.1.5. *Let V be a vector space. Then every basis of V has the same number of elements, or are all infinite.*

Proof. Let $B_1, B_2 \subset V$ be bases for V. Suppose that B_1 has infinitely many elements. Assume for a contradiction that B_2 has finitely many elements, say $n \in \mathbb{N}$. Take any subset $A_1 \subset B_1$ consisting of n vectors. Since A_1 is linearly independent and span $B_2 = V$, then by part (a) of the Replacement Lemma, we have that span $A_1 = V$. Then any $x \in B_1\backslash A_1$ can be written as a linear combination of the vectors in A_1. This implies that the set $A_1 \cup \{x\}$ is linearly dependent, which contradicts the fact that it is contained in the linearly independent set B_1. We therefore must have that B_2 is also an infinite set. This means that if V has at least one basis with infinitely many elements, then every basis of V also has infinitely many elements.

Now suppose that both B_1 and B_2 are finite bases for V with m and n elements respectively. Considering B_1 as a linearly independent set and B_2 as a spanning set for V, we obtain from part (b) of the Replacement Lemma that $m \leq n$. Reciprocally, considering B_2 as the linearly independent set and B_1 as the spanning set, we obtain $n \leq m$. This shows that $m = n$, so that every finite basis of V have the same number of elements. $\square$

This justifies the following definition:

Definition 1.1.6. Let V be a vector space. The **dimension** of V, denoted $\dim V$, is the number of elements in any basis of V if they are finite, or ∞ otherwise.

We will only consider finite dimensional vector spaces, that is, vector spaces in which there exists a basis with finitely many elements. Doing calculus on infinite dimensional vector spaces crosses into the realm of functional analysis, but we won't stray in that direction in the scope of these notes.

Given a basis $\{v_1, \ldots, v_n\}$ for a finite dimensional vector space V, the representation of a vector $v \in V$ in this particular basis is *unique*. That is, if $v = a_1 v_1 + \cdots + a_n v_n$ and $v = b_1 v_1 + \cdots + b_n v_n$ for constants $a_i, b_i \in \mathbb{R}$ for $i = 1, \ldots, n$, we have that $a_i = b_i$ for each i. To see this, take the difference of the two expressions

$$0 = (a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n.$$

By linear independence of the basis vectors, we conclude that $a_i - b_i = 0$ for each i, and the result follows.

One key advantage of working in finite dimensional spaces is that things all turn out to be significantly simpler. One easy consequence of the previous theorem is that in an n-dimensional vector space, if we have a linearly independent set consisting of n elements, it must actually span the entire space and is thus automatically a basis.

Corollary 1.1.7. *Let V be an n-dimensional vector space and let $\{v_1, \ldots, v_n\}$ be a linearly independent set. Then $\{v_1, \ldots, v_n\}$ is a basis for V.*

Proof. Suppose for a contradiction that $\operatorname{span}\{v_1, \ldots, v_n\} \neq V$. So there exists a $x \in V$ such that x is not a linear combination of the v_i's. Hence, $\{v_1, \ldots, v_n, x\}$ is a linearly independent set. So we have a linearly independent set consisting of $n + 1$ elements in an n-dimensional space. This is impossible, since every basis of V must contain n elements by Theorem 1.1.5 and no linearly independent set can have more elements than any basis of V by the Replacement Lemma. We must conclude that $V = \operatorname{span}\{v_1, \ldots, v_n\}$, and hence $\{v_1, \ldots, v_n\}$ is a basis for V. $\square$

Similarly, in an n-dimensional vector space, if we have a subset of V consisting of n elements that spans V, it must be linearly independent and hence a basis. The idea is that if it were linearly dependent, then the set contains 'redundant' information that can be removed without affecting the span. However, this would produce a spanning set with fewer elements than the dimension of the space - an impossibility.

Corollary 1.1.8. *Let V be an n-dimensional vector space and suppose $\{v_1, \ldots, v_n\} \subset V$ spans V. Then $\{v_1, \ldots, v_n\}$ is a basis for V.*

Proof. Suppose for a contradiction that $\{v_1, \ldots, v_n\}$ is a linearly dependent set. Then there exist constants $a_1, \ldots, a_n \in \mathbb{R}$ not all zero such that

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Without loss of generality, suppose that $a_n \neq 0$, so that

$$v_n = -\frac{1}{a_n}(a_1 v_1 + \cdots + a_{n-1} v_{n-1}).$$

Thus, $v_n \in \operatorname{span}\{v_1, \ldots, v_{n-1}\}$ and hence

$$\operatorname{span}\{v_1, \ldots, v_{n-1}\} = \operatorname{span}\{v_1, \ldots, v_n\} = V$$

i.e. we have found a spanning set for V consisting of $n - 1$ elements.

Let B be a basis for V. Since $\dim V = n$, then B is a linearly independent set consisting of n elements. Thus, by part (b) of the Replacement Lemma, we have that $n \leq n - 1$, which is absurd. Thus, we must have that $\{v_1, \ldots, v_n\}$ is linearly independent, and hence a basis for V. $\square$

Another simplification in the finite dimensional setting is that there actually aren't 'that many' different types of finite dimensional vector spaces for each dimension $n \in \mathbb{N}$ to study; we can easily completely classify every real finite dimensional vector space up to isomorphism - even better: they are all structurally identical to $\mathbb{R}^n$!

Theorem 1.1.9. *Let V be an n-dimensional vector space. Then V is isomorphic to $\mathbb{R}^n$.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a basis for V. Define the mapping $\phi : V \to \mathbb{R}^n$ to act on $v = a_1 v_1 + \cdots + a_n v_n$ by

$$\phi(v) = \phi(a_1 v_1 + \cdots + a_n v_n)$$
$$= (a_1, \ldots, a_n)^T.$$

So ϕ is the canonical map that maps a vector $v \in V$ to an n-tuple of real numbers, which are simply the coordinates of v with respect to a certain basis of V. We will show that ϕ is an isomorphism.

Note that the elements mapped by ϕ to $0 \in \mathbb{R}^n$ must have zero coefficients when expanded in terms of the basis $\{v_1, \ldots, v_n\}$, and thus can only be the zero vector. This shows that $\ker \phi = \{0\}$ and hence that ϕ is injective.

Let $x \in \mathbb{R}^n$. Then $x = (a_1, \ldots, a_n)^T$ for some $a_1, \ldots, a_n \in \mathbb{R}$. Thus, it is easy to see that the element $v = a_1 v_1 + \cdots + a_n v_n \in V$ is mapped to x by ϕ. Since the choice of $x \in \mathbb{R}^n$ was arbitrary, it follows that ϕ is surjective.

Let $v = a_1 v_1 + \cdots + a_n v_n \in V$ and $w = b_1 v_1 + \cdots + b_n v_n \in V$. Then

$$\phi(v + w) = (a_1 + b_1, \ldots, a_n + b_n)$$
$$= (a_1, \ldots, a_n) + (b_1, \ldots, b_n)$$
$$= \phi(v) + \phi(w).$$

Let $v \in V$ be as above and let $k \in \mathbb{R}$. Then

$$\phi(kv) = (ka_1, \ldots, ka_n)$$
$$= k(a_1, \ldots, a_n)$$
$$= k\phi(v).$$

Thus, ϕ is a vector space homomorphism. Together with the fact that ϕ is bijective, we have hence shown that ϕ is an isomorphism. $\square$

So any two n-dimensional vector spaces are isomorphic. This is what we meant when we said that the only n-dimensional vector space we will care about is $\mathbb{R}^n$ - all other instances of an n-dimensional vector space are algebraically equivalent, so we may as well just consider this simple example as being *the* n-dimensional vector space.

One more tidbit that we should address is to show that two vector spaces of different dimensions cannot be isomorphic, so that finite dimensional vector spaces are isomorphic if and only if their dimensions are equal (i.e. dimension is a vector space property that is preserved by isomorphism). This fact would imply that we have completely classified every single finite dimensional vector space: there is only one unique vector space for each dimension $n \in \mathbb{N}$ - namely $\mathbb{R}^n$, and $\mathbb{R}^n \simeq \mathbb{R}^m$ if and only if $m = n$.

Proposition 1.1.10. *Let V and W be finite dimensional vector spaces. If V and W are isomorphic, then $\dim V = \dim W$.*

Proof. Let $n = \dim V$ and let $\{v_1, \ldots, v_n\}$ be a basis for V. Let $\phi : V \to W$ be a vector space isomorphism. We will show that $\{\phi(v_1), \ldots, \phi(v_n)\}$ is a basis for W.

First, we prove linear independence. Consider the following vector equation:

$$a_1 \phi(v_1) + \cdots + a_n \phi(v_n) = 0$$

where $a_i \in \mathbb{R}$ for $i = 1, \ldots, n$. By linearity of the mapping ϕ, this is equivalent to:

$$\phi(a_1 v_1 + \cdots + a_n v_n) = 0.$$

So $a_1 v_1 + \cdots + a_n v_n \in \ker \phi$. However, since ϕ is injective, $\ker \phi = \{0\}$. It follows that:

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

By linear independence of $\{v_1, \ldots, v_n\}$, it follows that $a_i = 0$ for all $i = 1, \ldots, n$. Hence, $\{\phi(v_1), \ldots, \phi(v_n)\}$ is linearly independent.

Let $w \in W$. Since ϕ is surjective, there exists a $v \in V$ such that $w = \phi(v)$. Express v in terms of the basis $\{v_1, \ldots, v_n\}$:

$$v = c_1 v_1 + \cdots + c_n v_n.$$

Then:

$$\begin{aligned}
w &= \phi(v) \\
&= \phi(c_1 v_1 + \cdots + c_n v_n) \\
&= c_1 \phi(v_1) + \cdots + c_n \phi(v_n) \qquad \text{(by linearity of } \phi\text{).}
\end{aligned}$$

Hence, $w \in \text{span}\{\phi(v_1), \ldots, \phi(v_n)\}$. Since the choice of $w \in W$ is arbitrary, it follows that $W = \text{span}\{\phi(v_1), \ldots, \phi(v_n)\}$. This shows that $\{\phi(v_1), \ldots, \phi(v_n)\}$ is a basis for W, which has n elements. By definition, this implies that $\dim W = n = \dim V$. $\qquad \square$

In summary, we have proven the following very important result about finite dimensional vector spaces:

Theorem 1.1.11. *Let V and W be finite dimensional vector spaces. Then V and W are isomorphic if and only if $\dim V = \dim W$.*

Exercises

1. Let $T : V \to W$ be a linear map between vector spaces.

(a) Show that $\ker T$ is a subspace of V.

(b) Show that $\mathcal{R}(T)$ is a subspace of W.

(c) Show that T is injective if and only if $\ker T = \{0\}$.

   Hint: First show that $T(x) = T(y)$ if and only if $T(x - y) = 0$, making use of the linearity of T.

2. Consider the set $P \subset \mathbb{R}^3$ defined by

$$P = \left\{ (x, y, z)^T \in \mathbb{R}^3 \mid 2x + y - z = 0 \right\}.$$

(a) Show that P is a subspace of $\mathbb{R}^3$.

(b) Find a basis for P, and hence state $\dim P$.

3. Determine whether the following set of vectors is a linearly independent subset of $\mathbb{R}^4$

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \\ -6 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -4 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ -5 \\ -1 \end{pmatrix} \right\}.$$

1.2 Euclidean Space

Now that we have formulated all of the algebraic structure of the multidimensional spaces we will be working in, we will need to endow some further structure onto these spaces in order to do calculus. Of great importance is the ability to measure distances between points/vectors in our space. At the heart of analysis/calculus is the notion of *limits* and *convergence*. We will also frequently want to find bounds and estimates for quantities that may very well be vector-valued. So what we ultimately want is to prescribe some notion of 'length' for vectors, which will in turn naturally define a distance between two vectors (simply take the length of the difference of the two vectors). This is captured by the concept of a vector *norm*.

Definition 1.2.1. Let V be a vector space. A norm on V is a function $\|-\| : V \to \mathbb{R}$ satisfying:

(a) Positivity: For all $x \in V$ $\|x\| \geq 0$. Furthermore, $\|x\| = 0$ if and only if $x = 0$.

(b) Homogeneity of degree 1: For all $x \in V$ and $a \in \mathbb{R}$, $\|ax\| = |a|\|x\|$.

(c) Triangle inequality: for all $x, y \in V$, $\|x + y\| \leq \|x\| + \|y\|$.

A vector space together with a norm is called a **normed vector space**.

Another useful concept from coordinate geometry is the notion of angles. This will lend us the ability to define properties such as orthogonality, and perform operations such as projecting a vector onto a subspace. This structure is given rise to by an *inner product*, which we define below:

Definition 1.2.2. Let V be a vector space. A (real) **inner product** on V is a function $\langle -, - \rangle : V \times V \to \mathbb{R}$ satisfying:

(a) Symmetry: for all $x, y \in V$, $\langle x, y \rangle = \langle y, x \rangle$.

(b) Distributivity: for all $x, y, z \in V$, $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

(c) Homogeneity of degree 1: For all $x, y \in V$ and $a \in \mathbb{R}$, $\langle ax, y \rangle = a \langle x, y \rangle$.

(d) Positivity: For all $x \in V$, $\langle x, x \rangle \geq 0$. Moreover, $\langle x, x \rangle = 0$ if and only if $x = 0$.

A vector space together with an inner product is called an **inner product space**.

By symmetry of the inner product, we can deduce from the definition that $\langle x, ay \rangle = a \langle x, y \rangle$ and $\langle x, y + z \rangle = \langle x, z \rangle + \langle y, z \rangle$. Hence, an inner product on a vector space is a bilinear function (i.e. it becomes a linear function of one variable if we fix the value of the other input variable).

You are likely already quite familiar with an inner product on $\mathbb{R}^n$ from coordinate geometry, more commonly known as the dot product or scalar product. Given elements $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ and $y = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$, we define the dot product as

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

The notation $x \cdot y$ is more commonly used instead of $\langle x, y \rangle$ in this context. It is not too hard to see that the dot product is indeed an inner product on $\mathbb{R}^n$. The dot product of two vectors has the geometric interpretation as a value that quantifies how much the vectors point in the same direction: its magnitude is maximised when they are coincident and minimised (in fact, zero) when they are perpendicular.

One powerful fact about inner products is that they can be used to define a norm in a very natural way. Let V be an inner product space. Then for any $x \in V$, we set the norm of x to be given by $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$. We say that the inner product 'induces' a norm on V. So in some sense, once we had defined the notion of angles on our vector space, we have already implicitly fixed some compatible notion of distance as well for free. The positivity and homogeneity properties of the norm follow readily from the properties of the inner product and from the definition of the induced norm. Proving the triangle inequality will take a little more work, and will rely on an identity known as the *Schwarz inequality*:

Proposition 1.2.3. *(Schwarz Inequality) Let V be an inner product space. For every $x, y \in V$, we have*

$$\langle x, y \rangle \leq \|x\| \|y\|.$$

Proof. Let $x, y \in V$ and let $t \in \mathbb{R}$. By positivity of the inner product, we have that $\langle x - ty, x - ty \rangle \geq 0$. This can be rewritten as

$$\begin{aligned}
0 &\leq \langle x - ty, x - ty \rangle \\
&= \langle x, x \rangle - t \langle x, y \rangle - t \langle y, x \rangle + t^2 \langle y, y \rangle \\
&= \|x\|^2 - 2t \langle x, y \rangle + t^2 \|y\|^2.
\end{aligned}$$

So we have a real quadratic equation in the variable t. The quadratic can have at most a single zero, since $\langle x - ty, x - ty \rangle = 0$ if and only if $x - ty = 0$, i.e. $x = ty$. Hence, the discriminant of the quadratic must be either 0 (which corresponds to the single zero) or negative (no real zeros). The discriminant of the quadratic is

$$4 \langle x, y \rangle^2 - 4\|x\|^2 \|y\|^2.$$

Putting this all together, we conclude that

$$\begin{aligned}
4 \langle x, y \rangle^2 - 4\|x\|^2 \|y\|^2 &\leq 0 \\
\langle x, y \rangle &\leq \|x\| \|y\|.
\end{aligned}$$

$\square$

From the above proof, we note that equality occurs in the Schwarz inequality if and only if $x = ty$ (i.e. the vectors are scalar multiples of each other).

We are now ready to tackle the proof that the inner product gives rise to a vector norm:

Theorem 1.2.4. *Let V be an inner product space. The function $\|-\| : V \to \mathbb{R}$ defined by $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$ is a norm on V.*

Proof. Positivity: Let $x \in V$. We have that $\|x\|^2 = \langle x, x \rangle \geq 0$, and hence $\|x\| \geq 0$. Now, from $\|x\|^2 = \langle x, x \rangle$, we can easily see that $\|x\| = 0$ if and only if $x = 0$, by positivity of the inner product.

Homoegenity of degree 1: Let $x \in V$ and $a \in \mathbb{R}$. Then $\|ax\|^2 = \langle ax, ax \rangle = a^2 \langle x, x \rangle = a^2 \|x\|^2$. Taking square roots gives $\|ax\| = |a| \|x\|$.

Triangle inequality: Let $x, y \in V$. Then:

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y \rangle \\
&= \langle x, x \rangle + 2 \langle x, y \rangle + \langle y, y \rangle \\
&= \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \qquad \text{(by the Schwarz inequality)} \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Taking square roots gives the desired result: $\|x + y\| \leq \|x\| + \|y\|$. $\square$

Finally, we need to formalise the notion of measuring distance in our vector space. The idea of distance is captured by a 'metric', which intuitively speaking is a distance-measuring function.

Definition 1.2.5. Let X be a set. A **metric** on X is a function $d : X \times X \to \mathbb{R}$ satisfying:

(a) Positivity: for all $x, y \in X$, $d(x, y) \geq 0$. Moreover, $d(x, y) = 0$ if and only if $x = y$.

(b) Symmetry: for all $x, y \in X$, $d(x, y) = d(y, x)$.

(c) Triangle inequality: for all $x, y, z \in X$, $d(x, z) \leq d(x, y) + d(y, z)$.

A set together with a metric is called a **metric space**.

We mentioned earlier that once we had a norm on a vector space, this was sufficient to define a notion of distance. Specifically, we claimed that to measure the distance between two vectors x and y, we simply need to compute the length of the vector $x - y$ using the vector norm. Thus, we claim that the norm naturally induces a metric on the vector space.

Theorem 1.2.6. *Let V be a normed vector space. The function $d : V \times V \to \mathbb{R}$ defined by $d(x, y) = \|x - y\|$ is a metric on V.*

Proof. Positivity: Let $x, y \in V$. Then $d(x, y) = \|x - y\| \geq 0$ by positivity of the norm. Furthermore, $d(x, y) = 0$ if and only if $x - y = 0$, i.e. $x = y$.

Symmetry: Let $x, y \in V$. Then $d(x, y) = \|x - y\| = \|-(y - x)\| = |-1|\|y - x\| = d(y, x)$, which follows from the homogeneity of the norm.

Triangle inequality: Let $x, y, z \in V$. Then:

$$
\begin{aligned}
d(x, z) &= \|x - z\| \\
&= \|(x - y) + (y - z)\| \\
&\leq \|x - y\| + \|y - z\| \qquad \text{(since the norm obeys the triangle inequality)} \\
&= d(x, y) + d(y, z).
\end{aligned}
$$

$\square$

In summary, by defining an inner product on our vector space, the inner product induces a norm, which in turn induces a metric. So we get the geometric concepts of angle, length and distance all at once via an inner product. Hence, we call an n-dimensional vector space with an inner product **Euclidean n-space**; where 'Euclidean' refers to the fact that we have a means to measure distances in our space.

Returning to the example of $\mathbb{R}^n$, we see that $\mathbb{R}^n$ together with the dot product defines a Euclidean n-space. In fact, it is interesting to note that the metric induced by the dot product is simply the function that returns the Euclidean distance between two points in $\mathbb{R}^n$ (i.e. the Euclidean metric d_2):

$$
d_2(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.
$$

In the previous section, we showed that all n-dimensional vector spaces are isomorphic to each other, and in particular, $\mathbb{R}^n$, so we can just talk about $\mathbb{R}^n$ without any loss of generality. Even better, it is the case that all n-dimensional inner product spaces are isomorphic, and moreover the inner product is preserved by the isomorphism. Thus, there is essentially only one Euclidean n-space up to isomorphism: namely, $\mathbb{R}^n$ with the dot product. The proof of this will be the subject of the following section.

1.3 Orthonormal Basis

Most of the time when we are working with vector spaces, it is very useful to pick a basis and represent all vectors in terms of that basis. Some bases are more convenient to work with than others. Indeed, in a Euclidean space, now that we have some notion of angles provided by the inner product, we can define what it means for two vectors to be orthogonal (the generalisation of perpendicularity from coordinate geometry). Hence, we could try to find bases that are *orthonormal*; the basis vectors are orthogonal and normalised. We'll define these concepts below:

Definition 1.3.1. Let E be a Euclidean n-space. Two elements $x, y \in E$ are **orthogonal** if $\langle x, y \rangle = 0$.

A subset $A \subset E$ is said to be orthogonal if every pair of distinct elements from A are orthogonal.

Definition 1.3.2. Let E be a Euclidean n-space. A set $A \subset E$ is **orthonormal** if A is orthogonal and $\|x\| = 1$ for all $x \in A$.

Orthonormality is a powerful condition. One immediate consequence of the orthonormality of a set is linear independence:

Proposition 1.3.3. *Every orthonormal set in a Euclidean n-space is linearly independent.*

Proof. Suppose $\{e_1, \ldots, e_m\}$ is an orthonormal set in a Euclidean n-space E ($m \leq n$). Consider the vector equation $a_1 e_1 + \cdots + a_m e_m = 0$ for constants $a_1, \ldots, a_n \in \mathbb{R}$. Taking the inner product of e_i, $i = 1, \ldots, m$ with the above equation yields:

$$
\begin{aligned}
0 &= \left\langle e_i, \sum_{j=1}^{m} a_j e_j \right\rangle \\
&= \sum_{j=1}^{m} a_j \langle e_i, e_j \rangle \\
&= \sum_{j=1}^{m} a_j \delta_{ij} \\
&= a_i
\end{aligned}
$$

where we have made use of the Kronecker delta symbol δ_{ij}, which returns 1 if $i = j$ and 0 otherwise.

So $a_i = 0$ for all $i = 1, \ldots, m$. This proves the linear independence of the finite set $\{e_1, \ldots, e_m\}$.

Now suppose for a contradiction there exists an orthonormal set $A \subset E$ with more than n elements. Pick n elements from A to form the orthonormal set $\{e_1, \ldots, e_n\}$, which is linearly independent from what we have just proven. Hence, it forms a basis for E. Pick another *distinct* element $e_{n+1} \in A$, and write it as a linear combination of the basis elements:

$$
e_{n+1} = k_1 e_1 + \cdots + k_n e_n.
$$

Now take the inner product of the above equation with e_{n+1}:

$$
\begin{aligned}
\langle e_{n+1}, e_{n+1} \rangle &= \left\langle e_{n+1}, \sum_{j=1}^{n} k_j e_j \right\rangle \\
&= \sum_{j=1}^{n} k_j \langle e_{n+1}, e_j \rangle \\
&= 0.
\end{aligned}
$$

However, $\langle e_{n+1}, e_{n+1} \rangle = \|e_{n+1}\|^2 = 1$, so we have that $1 = 0$; a contradiction. It follows that there cannot exist an orthonormal set in E with more than n elements. Since we have proven that every orthonormal set with $m \leq n$ elements is linearly independent, then we thus conclude that every orthonormal set in E is linearly independent. $\qquad\square$

One interesting consequence of the above proof is the following. Suppose we could write a vector x from a Euclidean space E as a linear combination of orthonormal vectors $x = a_1 e_1 + \cdots + a_n e_n$. Then we can define a mapping $\pi_i : E \to \mathbb{R}$ for $i = 1, \ldots, n$ by $\pi_i(x) = \langle x, e_i \rangle = a_i$. This illustrates another advantage of using an orthonormal basis to represent vectors: finding the coefficients becomes a straightforward task - one simply needs to calculate $\pi_i(x)$ to find the ith coefficient. We call the mapping π_i a *projection operator*; essentially π_i is projecting the vector x onto the subspace spanned by the ith basis vector and picking out the amplitude of the vector projection.

Now that we are sufficiently hyped up about orthonormal bases, the next natural question to ask is whether it is always possible to find an orthonormal basis for any Euclidean n-space. We know from the

previous result that every orthonormal set is linearly independent. Combined with the fact that in an n-dimensional space, any linearly independent set consisting of n elements is automatically a basis, all we need to do is prove the existence of an orthonormal set consisting of n elements. Spoiler alert: yes, this is certainly possible. Better yet, the proof of this result will be constructive, it illustrates a standard algorithm called the 'Gram-Schmidt' process which transforms an arbitrary basis into an orthonormal basis.

Proposition 1.3.4. *Let E be a Euclidean n-space. Then there exists an orthonormal subset $A \subset E$ consisting of n elements.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a linearly independent set in E. It is worth pointing out that none of the vectors are 0 due to linear independence. Thus, we can create an orthonormal set of one element by taking the first element from the set and normalising it: define $e_1 = v_1 / \|v_1\|$. Then $\{e_1\}$ is an orthonormal set, and e.g. $\{e_1, v_2\}$ is a linearly independent set (since e_1 is simply a scalar multiple of v_1). This sets us up to prove the inductive step.

Suppose that we had that $\{e_1, \ldots, e_k\}$ is an orthonormal set ($k < n$) and that $\{e_1, \ldots, e_k, v_{k+1}\}$ is a linearly independent set. Define:

$$y_{k+1} = v_{k+1} - \sum_{j=1}^{k} \langle v_{k+1}, e_j \rangle \, e_j.$$

It follows that $y_{k+1} \neq 0$ since y_{k+1} is a linear combination of $e_1, \ldots, e_k, v_{k+1}$, which is a linearly independent set - and so the only way for such a linear combination to yield the zero vector is if all of the coefficients are 0, which is not the case here. Hence, $\|y_{k+1}\| \neq 0$, and so we can safely define:

$$e_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|}.$$

So $\|e_{k+1}\| = 1$, and for all $i = 1, \ldots, k$:

$$
\begin{aligned}
\langle e_{k+1}, e_i \rangle &= \left\langle \frac{1}{\|y_{k+1}\|} \left(v_{k+1} - \sum_{j=1}^{k} \langle v_{k+1}, e_j \rangle \, e_j \right), e_i \right\rangle \\
&= \frac{1}{\|y_{k+1}\|} \left(\langle v_{k+1}, e_i \rangle - \sum_{j=1}^{k} \langle v_{k+1}, e_i \rangle \, \langle e_j, e_i \rangle \right) \\
&= \frac{1}{\|y_{k+1}\|} \left(\langle v_{k+1}, e_i \rangle - \sum_{j=1}^{k} \langle v_{k+1}, e_j \rangle \, \delta_{ij} \right) \\
&= \frac{1}{\|y_{k+1}\|} \left(\langle v_{k+1}, e_i \rangle - \langle v_{k+1}, e_i \rangle \right) \\
&= 0.
\end{aligned}
$$

Hence, $\{e_1, \ldots, e_{k+1}\}$ is an orthonormal set. If $k+1 < n$, then there is still an element v_{k+2} that cannot be written as a linear combination of the orthonormal set $\{e_1, \ldots, e_{k+1}\}$, since the e_i's are simply linear combinations of $v_1, \ldots, v_{k+1}$, but $\{v_1, \ldots, v_{k+2}\}$ is a linearly independent set. It follows that $\{e_1, \ldots, e_{k+1}, v_{k+2}\}$ is a linearly independent set. We can then proceed via induction to deduce that we can produce an orthonormal set $\{e_1, \ldots, e_n\}$. $\square$

The previous two results together imply the following:

Theorem 1.3.5. *Every Euclidean n-space has an orthonormal basis.*

Let us return to the original problem that we wished to address in this section: that there is only one Euclidean n-space up to isomorphism. Since we have already proven that the vector space structure of two Euclidean n-spaces are equivalent, we simply need to prove that the 'extra structure' of the Euclidean space is also preserved under a vector space isomorphism. Let's first make explicit what 'extra stuff' needs to be preserved:

Definition 1.3.6. Let E and F be Euclidean spaces. We say that E and F are **isomorphic** if there exists a mapping $\phi : E \to F$ such that ϕ is a vector space isomorphism that also preserves the inner product, i.e. for all $x, y \in E$, $\langle x, y \rangle = \langle \phi(x), \phi(y) \rangle$.

So we require that just the inner product needs to be preserved by the isomorphism. But what about the corresponding norm and metric? Well, if the inner product between Euclidean spaces E and F is preserved by an isomorphism $\phi : E \to F$, it follows that ϕ also preserves the induced norms and metrics on these spaces. To see this, let $x \in E$ and notice that $\|x\|^2 = \langle x, x \rangle = \langle \phi(x), \phi(x) \rangle = \|\phi(x)\|^2$. Now that we have shown that ϕ is norm-preserving, the fact that distances are preserved follows easily. For any $x, y \in E$ we have that $d(x, y) = \|x - y\| = \|\phi(x - y)\| = \|\phi(x) - \phi(y)\| = d(\phi(x), \phi(y))$. Hence, we see that ϕ is an isometry (a distance preserving function).

Recall from our proof that every n-dimensional vector space is isomorphic (to $\mathbb{R}^n$) (Theorem 1.1.9), we constructed a function that essentially mapped one basis onto another. In order to preserve the inner product of the Euclidean spaces, the isomorphisms that we will construct in the following proof will be a subset of the isomorphisms that did the trick in the more general vector space case. Specifically, the isomorphisms between Euclidean spaces will, loosely speaking, send *orthonormal bases to orthonormal bases*.

Theorem 1.3.7. *Let E and F be Euclidean n-spaces. Then E and F are isomorphic.*

Proof. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis in E, and $\{\overline{e_1}, \ldots, \overline{e_n}\}$ an orthonormal basis in F (we know that orthonormal bases for Euclidean spaces exist by the previous theorem which we have just proven). We can express any $x \in E$ as a linear combination of the basis vectors: $x = a_1 e_1 + \cdots + a_n e_n$. We can hence define the map $\phi : E \to F$ by

$$\phi(x) = a_1 \overline{e_1} + \cdots + a_n \overline{e_n}.$$

The form of ϕ is exactly that of the isomorphism which we have constructed in the proof of Theorem 1.1.9, i.e. a map that sends a vector represented in a given basis to a vector in the other space with the same coordinates/coefficients but with respect to a basis in the target space[1]. Thus, we already have that ϕ is a vector space isomorphism. We thus simply need to demonstrate that ϕ preserves the inner product. Let $x \in E$ as above and let $y = b_1 e_1 + \cdots + b_n e_n \in E$. Then

$$
\begin{aligned}
\langle x, y \rangle &= \left\langle \sum_{i=1}^{n} a_i e_i, \sum_{j=1}^{n} b_j e_j \right\rangle \\
&= \sum_{i=1}^{n} a_i \left\langle e_i, \sum_{j=1}^{n} b_j e_j \right\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \langle e_i, e_j \rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \delta_{ij} \qquad \text{(by orthonormality of the basis vectors)} \\
&= \sum_{i=1}^{n} a_i b_i.
\end{aligned}
$$

---

[1]Specifically, we mapped a basis from an n-dimensional vector space to the standard basis in $\mathbb{R}^n$ in that proof.

But we also have that

$$\langle \phi(x), \phi(y) \rangle = \left\langle \sum_{i=1}^{n} a_i \overline{e_i}, \sum_{j=1}^{n} b_j \overline{e_j} \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \langle \overline{e_i}, \overline{e_j} \rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \delta_{ij} \qquad \text{(by orthonormality of the basis vectors)}$$

$$= \sum_{i=1}^{n} a_i b_i.$$

So $\langle x, y \rangle = \langle \phi(x), \phi(y) \rangle$, and so the inner product is preserved. Hence, ϕ is an isomorphism of Euclidean spaces. $\qquad \square$

Hence, every Euclidean n-space is 'the same' from a structural point of view, so we might as well work with the easiest example of it we can think of: i.e. $\mathbb{R}^n$ with the dot product.

While on the topic of orthonormal bases, we'll conclude this section with a fun digression. Let E be a Euclidean n-space, and suppose we had two orthonormal bases in E: $\{e_1, \ldots, e_n\}$ and $\{\overline{e_1}, \ldots, \overline{e_n}\}$. We can write the barred basis vectors $\overline{e_i}$ in terms of the basis vectors e_j:

$$\overline{e_i} = \sum_{j=1}^{n} a_{ij} e_j \quad i = 1, \ldots, n.$$

The n^2 constants $a_{ij} \in \mathbb{R}$ form a $n \times n$ matrix O, which represents the linear map $\phi : E \to E$ defined by $\phi(a_1 e_1 + \cdots + a_n e_n) = a_1 \overline{e_1} + \cdots + a_n \overline{e_n}$ (more on this in Section 1.10), which we know from above to be an automorphism of E. So if we wrote the components of a vector $x \in E$ with respect to the orthonormal basis $\{e_1, \ldots, e_n\}$ as a column vector, then $\phi(x) = Ox$, where the components of $\phi(x)$ are also expressed in a column vector. What properties does the matrix O have? Well, since $\langle \overline{e_i}, \overline{e_j} \rangle = \delta_{ij}$ by orthonormality, we can substitute our expression for the barred basis vectors in terms of the unbarred basis vectors to obtain:

$$\delta_{ij} = \left\langle \sum_{k=1}^{n} a_{ik} e_k, \sum_{l=1}^{n} a_{jl} e_l \right\rangle$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} a_{jl} \langle e_k, e_l \rangle$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} a_{jl} \delta_{kl}$$

$$= \sum_{k=1}^{n} a_{ik} a_{jk}.$$

Suppose that b_{ij} are the components of the matrix O^T (the matrix transpose of O). Then $b_{kj} = a_{jk}$, so that we have that $\delta_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$. Since δ_{ij} are the components of the $n \times n$ identity matrix I_n, we can write the above equation in the following matrix form

$$I_n = OO^T.$$

Since O is the matrix representation of a vector space automorphism, it is invertible, and hence by premultiplying both sides of the above equation by O^{-1}, we arrive at

$$O^{-1} = O^T$$

i.e. the matrix O is an orthogonal matrix. This reveals that orthogonal transformations on Euclidean spaces are simply the ones that map orthonormal bases to orthonormal bases.

1.4   Dual Space

In this section, we study a very special type of vector space mapping. Namely, the linear maps that take a vector as an input and returns a scalar. Such maps are called **linear functionals**.

Definition 1.4.1. Let V be an n-dimensional vector space. A **linear functional** on V is a linear mapping $f : V \to \mathbb{R}$.

Definition 1.4.2. Let V be an n-dimensional vector space. The **dual space** V^* is the set of all linear functionals on V.

The dual space of a vector space is itself a vector space under the operations of addition and scalar multiplication of real-valued functions. Elements of the dual space are sometimes also referred to as *covectors*.

There are many important linear functionals that are central to analysis. To name a few, the trace operator, the inner product with one input fixed, and integral operators. It is therefore quite valuable to study them in a general setting.

In the finite dimensional setting, the study of linear functionals and the dual space is quite simple. In the following proposition, we show that every linear functional on a finite dimensional vector space can be completely characterised simply by knowing what values it takes at finitely many points (specifically, at the basis vectors), and that we can essentially cook up a linear functional to behave in 'any way we want' (i.e. take specific values at certain desired points).

Proposition 1.4.3. *Let V be an n-dimensional vector space and let $\{v_1, \ldots, v_n\}$ be a basis for V. Then:*

(a) Every linear functional $f \in V^*$ is completely determined by its values at $v_1, \ldots, v_n$.

(b) For any real numbers $a_1, \ldots, a_n \in \mathbb{R}$ there exists an $f \in V^*$ such that $f(v_i) = a_i$ for all $i = 1, \ldots, n$.

Proof. (a) Let $f \in V^*$ and let $x = k_1 v_1 + \cdots + k_n v_n \in V$. Then

$$
\begin{aligned}
f(x) &= f(k_1 v_1 + \cdots + k_n v_n) \\
&= k_1 f(v_1) + \cdots + k_n f(v_n) \qquad\qquad \text{(by linearity of } f\text{)}.
\end{aligned}
$$

One sees from the above expression that $f(x)$ can be computed using only the knowledge of the values $f(v_1), \ldots, f(v_n)$. Moreover, this *uniquely* identifies the linear functional. Suppose $g \in V^*$ agrees with f at each of the basis vectors, i.e. $f(v_i) = g(v_i)$ for all $i = 1, \ldots, n$. Then for any $x \in V$ with representation as above, we have that

$$
\begin{aligned}
g(x) &= g(k_1 v_1 + \cdots + k_n v_n) \\
&= k_1 g(v_1) + \cdots + k_n g(v_n) \\
&= k_1 f(v_1) + \cdots + k_n f(v_n) \\
&= f(x).
\end{aligned}
$$

(b) Let $a_1, \ldots, a_n \in \mathbb{R}$. For $x = k_1 v_1 + \cdots + k_n v_n \in V$, define the mapping $f : V \to \mathbb{R}$ by

$$
f(x) = k_1 a_1 + \cdots + k_n a_n.
$$

Then $f(v_i) = a_i$ for all $i = 1, \ldots, n$, as required. We'll now set about to show that f is linear.

Let $x \in V$ as above and let $y = l_1 v_1 + \ldots l_n v_n \in V$. Then

$$
\begin{aligned}
f(x + y) &= (k_1 + l_1)a_1 + \cdots + (k_n + l_n)a_n \\
&= k_1 a_1 + \cdots + k_n a_n + l_1 a_1 + \cdots + l_n a_n \\
&= f(x) + f(y).
\end{aligned}
$$

Let $c \in \mathbb{R}$. Then

$$
\begin{aligned}
f(cx) &= ck_1v_1 + \cdots + ck_nv_n \\
&= c(k_1v_1 + \cdots + k_nv_n) \\
&= cf(x).
\end{aligned}
$$

This proves that f is a linear map, so that $f \in V^*$. $\qquad\square$

Given a vector space, how big is its dual space? In the infinite dimensional case, it can be shown via the axiom of choice that the dual space is always 'bigger' than the original space. In the finite dimensional case however, again we have it easy: the dimensional of the dual space is the same as that of original space, so that they are isomorphic.

Theorem 1.4.4. *Let V be an n-dimensional vector space. Then its dual space V^* is also an n-dimensional vector space.*

Proof. We will prove that V^* is isomorphic to $\mathbb{R}^n$ as vector spaces, which will imply that V^* has the same dimension as $\mathbb{R}^n$, which is n.

We first need to construct the function between V^* and $\mathbb{R}^n$ which will serve as the candidate for what we hope will be the isomorphism. Fix a basis $\{v_1, \ldots, v_n\}$ for V. Let $f \in V^*$. Define $\psi : V^* \to \mathbb{R}^n$ by

$$
\psi(f) = (f(v_1), \ldots, f(v_n))^T.
$$

Let $f, g \in V^*$. Then

$$
\begin{aligned}
\psi(f + g) &= (f(v_1) + g(v_1), \ldots, f(v_n) + g(v_n))^T \\
&= (f(v_1), \ldots, f(v_n))^T + (g(v_1), \ldots, g(v_n))^T \\
&= \psi(f) + \psi(g).
\end{aligned}
$$

Let $a \in \mathbb{R}$. Then

$$
\begin{aligned}
\psi(af) &= (af(v_1), \ldots, af(v_n))^T \\
&= a(f(v_1), \ldots, f(v_n))^T \\
&= a\psi(f).
\end{aligned}
$$

Thus, ψ is a linear map. We now proceed to show that ψ is a bijection.

Suppose $\psi(f) = \psi(g)$. Then $(f(v_1), \ldots, f(v_n))^T = (g(v_1), \ldots, g(v_n))^T$, i.e. f and g agree at each of the basis vectors. By part (a) Proposition 1.4.3, this implies that $f = g$, thus proving the injectivity of ψ.

Now let $x = (a_1, \ldots, a_n)^T \in \mathbb{R}^n$. By part (b) of Proposition 1.4.3, there exists an $f \in V^*$ such that $f(v_i) = a_i$ for all $i = 1, \ldots, n$. It follows that $\psi(f) = x$. Since the choice of $x \in \mathbb{R}^n$ was arbitrary, it follows that $\mathcal{R}(\psi) = \mathbb{R}^n$, thus proving surjectivity of ψ. Putting all of this together shows that ψ is a bijective linear function, i.e. a vector space isomorphism. $\qquad\square$

Let V be an n-dimensional vector space. Since the dual space V^* is n-dimensional, it has a basis consisting of n elements. Given a basis $\{v_1, \ldots, v_n\}$ for V, there is a particular nice choice for a basis on V^*. For each $i = 1, \ldots, n$, define the linear functional $v_i^* : V \to \mathbb{R}$ by

$$
v_i^*(v_j) = \delta_{ij}, \quad \text{for } j = 1, \ldots, n.
$$

So the linear functionals v_i^* satisfy a 'bi-orthogonality property'. The fact that such linear functionals exist again comes from Proposition 1.4.3. The set $\{v_1^*, \ldots, v_n^*\}$ is a linearly independent set in V^*. To see

this, let $x = a_1 v_1 + \cdots + a_n v_n \in V$ be arbitrary and consider the following vector equation

$$
\begin{aligned}
0 &= \sum_{i=1}^{n} c_i v_i^*(x) \\
&= \sum_{i=1}^{n} c_i v_i^* \left(\sum_{j=1}^{n} a_j v_j \right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i a_j v_i^*(v_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i a_j \delta_{ij} \\
&= \sum_{i=1}^{n} c_i a_i.
\end{aligned}
$$

Since the above equation must hold for all $x \in V$, choosing $x = v_k$ for each $k = 1, \ldots, n$ means that $a_i = \delta_{ik}$ in the above expression, which gives $c_k = 0$. This demonstrates the linear independence of the v_i^*'s. Since we have shown that V^* is an n-dimensional vector space, and we have that $\{v_1^*, \ldots, v_n^*\}$ is a linearly independent set in V^* consisting of n elements, we know automatically that this must be a basis for V^*. We call this special basis the **dual basis**.

Now, since the dual space V^* is itself a vector space, we can consider its dual space V^{**}, called the **second dual**. We know that V^{**} must also be a vector space with the same dimension as V, so they are isomorphic. Better yet, there is a natural identification of elements from V with elements from V^{**}, so that the isomorphism is canonical.

Theorem 1.4.5. *Let V be an n-dimensional vector space. Then V is canonically isomorphic to its second dual V^{**}.*

Proof. For each $x \in V$, define the mapping $\phi_x : V^* \to \mathbb{R}$ by

$$\phi_x(f) = f(x)$$

i.e. ϕ_x is an evaluation map. Given a linear functional on V^*, it returns a scalar by evaluating it at the fixed point x. We will show that ϕ_x is a linear map.

Let $f, g \in V^*$. Then $\phi_x(f + g) = (f + g)(x) = f(x) + g(x) = \phi_x(f) + \phi_x(g)$.
Let $a \in \mathbb{R}$. Then $\phi_x(af) = af(x) = a\phi_x(f)$.

Thus, ϕ_x is a linear functional on V^*, so $\phi_x \in V^{**}$. We can hence define the mapping $\psi : V \to V^{**}$ by $\psi(x) = \phi_x$. We will show that ψ is a vector space homomorphism. Let $x, y \in V$ and let $f \in V^*$. Then

$$
\begin{aligned}
(\psi(x + y))(f) &= \phi_{x+y}(f) \\
&= f(x + y) \\
&= f(x) + f(y) \\
&= \phi_x(f) + \phi_y(f) \\
&= (\psi(x))(f) + (\psi(y))(f).
\end{aligned}
$$

Since the choice of $f \in V^*$ was arbitrary, this shows that $\psi(x + y) = \psi(x) + \psi(y)$.

Let $a \in \mathbb{R}$. Then

$$
\begin{aligned}
(\psi(ax))(f) &= \phi_{ax}(f) \\
&= f(ax) \\
&= af(x) \\
&= a\phi_x(f) \\
&= a(\psi(x))(f).
\end{aligned}
$$

Thus, $\psi(ax) = a\psi(x)$. This proves that ψ is a vector space homomorphism.

To show the injectivity of ψ, we will prove that its kernel is trivial. Suppose $x \in \ker \psi$, so $\psi(x) = 0$. This means that for all $f \in V^*$, $f(x) = 0$. If $x \neq 0$, then with respect to some basis $\{v_1, \ldots, v_n\}$ of V, x can be written as a nontrivial linear combination of the basis vectors: $x = a_1 v_1 + \cdots + a_n v_n$, where at least one of the a_i's are nonzero. Pick a nonzero coefficient and call it a_k. Then by part (b) of Proposition 1.4.3, there exists a $g \in V^*$ such that $g(v_k) = 1$ and $g(v_i) = 0$ for all $i \neq k$. Then

$$
\begin{aligned}
g(x) &= g(a_1 v_1 + \cdots + a_n v_n) \\
&= a_1 g(v_1) + \cdots + a_k g(v_k) + \cdots + a_n g(v_n) \\
&= a_k \\
&\neq 0.
\end{aligned}
$$

This is a contradiction. Hence, $x = 0$, which shows that $\ker \psi = \{0\}$, and hence ψ is injective.

Finally, we show that ψ is surjective. Let $\chi \in V^{**}$. Let $\{v_1^*, \ldots, v_n^*\}$ be the dual basis for $\{v_1, \ldots, v_n\}$. Take $a_i = \chi(V_i^*)$ for $i = 1, \ldots, n$, and thus construct the element $x = a_1 v_1 + \cdots + a_n v_n \in V$. We claim that $\chi = \phi_x$. Let $f \in V^*$, and expand it in terms of the dual basis

$$
f = \sum_{i=1}^{n} c_i v_i^*.
$$

Then we have

$$
\begin{aligned}
\phi_x(f) &= f(x) \\
&= \sum_{i=1}^{n} c_i v_i^*(x) \\
&= \sum_{i=1}^{n} c_i v_i^* \left(\sum_{j=1}^{n} a_j v_j \right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i a_j v_i^*(v_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i a_j \delta_{ij} \\
&= \sum_{i=1}^{n} c_i a_i \\
&= \sum_{i=1}^{n} c_i \chi(v_i^*) \\
&= \chi \left(\sum_{i=1}^{n} c_i v_i^* \right) \qquad \text{(by linearity of } \chi) \\
&= \chi(f).
\end{aligned}
$$

The above argument holds for all $f \in V^*$, and thus we conclude that $\chi = \phi_x = \psi(x)$. This proves that ψ is surjective, and hence a bijection, and hence an isomorphism. $\qquad\square$

The property that a vector space is naturally isomorphic to its second dual is referred to as *algebraic reflexivity*. This is another property that is guaranteed to hold in the finite dimensional setting, but not in the infinite dimensional case. There is always a natural monomorphism (injective homomorphism) from a vector space to its second dual, however it is an isomorphism if and only if the vector space is finite dimensional.

Naturally, we can keep playing the game of successively finding the dual space of a dual space, but the results in this section show that this is not necessary. It turns out in the finite dimensional setting that a vector space and its dual are of equal dimension, and are isomorphic. So the study of the dual space of a finite dimensional vector space reduces to the study of the vector space itself.

1.5   Topology of Euclidean Space

So far, we have mostly concerned ourselves with the algebraic properties of Euclidean space. We will now consider the structure of Euclidean space that makes it amenable to do calculus on. Recall that the Euclidean inner product induces a vector norm, which in turn induces a metric. To add to this already lengthy chain of implications, the metric induces a *topology* on our Euclidean space. A topological space is the most general setting in which one can define the notion of continuity, which is certainly something we would like to have in a calculus setting!

Since the topology of Euclidean space is the metric topology (where the metric is the Euclidean metric d_2), we will define the open sets via the notion of *open balls*. We have already described the Euclidean metric for $\mathbb{R}^n$. To generalise it to Euclidean n-space E, take an orthonormal basis $\{e_1, \ldots, e_n\}$ of E. Then for any $x = x_1 e_1 + \cdots + x_n e_n \in E$, the Euclidean norm is given by

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$$

with corresponding Euclidean metric

$$d_2(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

where $y = y_1 e_1 + \cdots + y_n e_n$.

Definition 1.5.1. Let E be a Euclidean n-space. An **open ball** of centre $x \in E$ and radius $r > 0$ is the set $B(x, r) = \{y \in E \mid d_2(x, y) < r\} = \{y \in E \mid \|x - y\| < r\}$.

The **closed ball** of centre x and radius r is the set $\overline{B}(x, r) = \{y \in E \mid d_2(x, y) \leq r\} = \{y \in E \mid \|x - y\| \leq r\}$.

An open ball centered at a point $x \in E$ with radius r simply consists of all of the points that are within a distance r of x. The closed ball also includes those points that are exactly a distance of r away from x (as measured using the Euclidean metric).

Definition 1.5.2. Let E be a Euclidean n-space. A set $U \subset E$ is said to be **open** if for every $x \in U$, there exists an $r > 0$ such that $B(x, r) \subset U$.

Some trivial examples of open sets include the empty set, and the whole space E. Open balls themselves are also open sets. An open rectangle, which is simply the Cartesian product of open intervals in $\mathbb{R}$, is an open set in $\mathbb{R}^n$.

Now we address the age-old question of: 'once we have a bunch of open sets, how do we make new open sets from old?' This is answered by:

Proposition 1.5.3. *Let E be a Euclidean n-space. Then:*

(a) If $U_\alpha \subset E$ is open for each $\alpha \in A$, then the union

$$U = \bigcup_{\alpha \in A} U_\alpha$$

is open in E.

(b) If $U_1, \ldots, U_k$ are open subsets of E, then the intersection

$$U = U_1 \cap \cdots \cap U_k$$

is open in E.

Proof. (a) Let $U_\alpha \subset E$ be open for each $\alpha \in A$, and consider the set $U = \bigcup_{\alpha \in A} U_\alpha$. Let $x \in U$. Then there exists a $\beta \in A$ such that $x \in U_\beta$. Since U_β is open, there exists an $r > 0$ such that $B(x, r) \subset U_\beta \subset U$. Thus, U is open.

(b) Let $U_1, \ldots, U_k$ be open, and consider the set $U = U_1 \cap \cdots \cap U_k$. Let $x \in U$. For each $i = 1, \ldots, k$, $x \in U_i$, so there exists an $r_i > 0$ such that $B(x, r_i) \subset U_i$. Choose $r = \min\{r_1, \ldots, r_k\} > 0$. Then $B(x, r) \subset U_i$ for all $i = 1, \ldots, k$. Thus, $B(x, r) \subset U$, and hence U is open. $\square$

So arbitrary unions of open sets are open, and *finite* intersections of open sets are open. Infinite intersections of open sets are not necessarily open. For example, consider $I_n = \left(-\frac{1}{n}, \frac{1}{n}\right)$, which is an open set in $\mathbb{R}$ for each $n \in \mathbb{N}$. But

$$\bigcap_{n=1}^{\infty} I_n = \{0\}$$

which is *not* open in $\mathbb{R}$.

Proposition 1.5.3 above, in combination with the fact that $\emptyset$ and E are open, shows that the collection of open sets in a Euclidean space E forms a *topology* on E. This fact did not appeal to anything particular about Euclidean space other than its underlying metric space structure.

We will later want to define the notions of limits and continuity, which both involve the idea of points being forced to being 'really close to each other'. Frequently in analysis we will also allow for our variables to vary in a 'sufficiently small amount'. To make formal these ideas, we use the topologcal concept of 'closeness' in terms of *neighbourhoods* of points:

Definition 1.5.4. Let E be a Euclidean n-space and let $x \in E$. A **neighbourhood** of x is a subset $U \subset E$ such that there exists an open subset $V \subset U$ that contains x.

Note that in the metric topology, an open set containing a point $x \in E$ must contain an open ball $B(x, r)$ for some $r > 0$, so that we can also state our definition of a neighbourhood of a point x as simply 'a set which contains an open ball centred at x'. Hence, the topological notion of 'wiggling around x in an arbitrarily small amount' (constraining variation to smaller and smaller neighbourhoods) equates to the metric notion of small movement (constraining variation to smaller and smaller open balls).

Thus, an alternative way to characterise open sets is the statement that 'a set is open if and only if it is a neighbourhood of all of its points', which follows immediately from the definition. We can also rephrase this in terms of *interior points*, which we define below:

Definition 1.5.5. Let E be a Euclidean n-space. An **interior point** of a set $S \subset E$ is an $x \in S$ such that there exists an $r > 0$ such that $B(x, r) \subset S$.

One sees immediately from the definition above and from the definition of an open set that a set is open if and only if it all of its points are interior points. So if a set S has any points that are not interior points, it cannot be open. However, if we consider the subset of S comprising of only the interior points of S, then we get an open subset of S. In fact, this is the largest open subset contained in S, which we call the *interior* of S.

Definition 1.5.6. Let E be a Euclidean n-space and let $S \subset E$. The **interior** of S, denoted S°, is the union of all open subsets of S.

Our definition of the interior above formulates it as the largest open subset contained in S, but it can be readily shown that the interior is simply the set of interior points of S. This means that a set is open if and only if it is equal to its interior.

Also equally important in a topological space is the notion of a *closed* set:

Definition 1.5.7. Let E be a Euclidean n-space. A set $C \subset E$ is said to be **closed** if it is the complement of an open set, i.e. $C = E \backslash U$ for some open set $U \subset E$.

One should be careful that a set being closed is *not* the negation of a set being open. There are sets that are neither open nor closed (e.g. the half-open interval in $\mathbb{R}$: $[a, b)$), and sets that are *both* open and closed (the empty set and the entire space).

Closed balls are a simple example of closed sets. Cartesian products of closed intervals in $\mathbb{R}$ are closed sets in $\mathbb{R}^n$. An important example of closed sets in Euclidean space are vector subspaces. Complementary to Proposition 1.5.3, we now list the ways one can combine closed sets to make new closed sets:

Proposition 1.5.8. *Let E be a Euclidean n-space. Then:*

(a) *If $C_\alpha \subset E$ is closed for each $\alpha \in A$, then the intersection*

$$C = \bigcap_{\alpha \in A} C_\alpha$$

is closed in E.

(b) *If $C_1, \ldots, C_k$ are closed subsets of E, then the union*

$$C = C_1 \cup \cdots \cup C_k$$

is closed in E.

Proof. (a) Let $C_\alpha \subset E$ be closed for each $\alpha \in A$. Then $E \backslash C_\alpha$ is open for all $\alpha \in A$. Hence, $\bigcup_{\alpha \in A}(E \backslash C_\alpha)$ is an open set, since arbitrary unions of open sets are open. But by de Morgan's Laws, we have that

$$\bigcup_{\alpha \in A}(E \backslash C_\alpha) = E \backslash \left(\bigcap_{\alpha \in A} C_\alpha \right).$$

We thus see that $C = \bigcap_{\alpha \in A} C_\alpha$ is closed.

(b) Let $C_1, \ldots, C_k$ be closed. Then $E \backslash C_i$ is open for all $i = 1, \ldots, k$. Thus, $(E \backslash C_1) \cap \cdots \cap (E \backslash C_k)$ is open since finite intersections of open sets are open. But by de Morgan's Laws, we have that

$$(E \backslash C_1) \cap \cdots \cap (E \backslash C_k) = E \backslash (C_1 \cup \cdots \cup C_k).$$

It immediately follows that $C_1 \cup \cdots \cup C_k$ is closed. $\square$

One way to study closed sets, and the degree at which a set fails to be closed is via the notion of *limit points*:

Definition 1.5.9. Let E be a Euclidean n-space and let $S \subset E$. A **limit point** of S is a point $x \in E$ such that for every $r > 0$, $S \cap (B(x, r) \backslash \{x\}) \neq \emptyset$.

The motivation for the name 'limit point' arises from the fact you can approach a limit point of a set and get arbitrarily close to it even whilst remaining completely within the set. That is, it is the limit of a nonconstant sequence consisting entirely of points from the set.

One important fact about closed sets is that the limit of a sequence consisting of points from the set must stay inside the set. From the above discussion, it is then clear that a closed set must contain all of its limit points. This turns out to be both a necessary and sufficient criteria for a set to be closed.

Proposition 1.5.10. *Let E be a Euclidean n-space. A set is closed in E if and only if it contains all of its limit points.*

Proof. ($\implies$) Suppose $S \subset E$ is closed, but assume for a contradiction that S does not contain all of its limit points. Let $x \in E$ be a limit point of S but $x \notin S$. Then $x \in E \backslash S$. However, we know that $E \backslash S$ is an open set (since S is closed), so there exists an $r > 0$ such that $B(x, r) \subset E \backslash S$. But since x is a limit point of S, then there exists a $y \in S \cap (B(x, r) \backslash \{x\})$, so that $y \in S \cap E \backslash S$, which is impossible. Thus, S must contain all of its limit points.

($\impliedby$) Suppose S contains all of its limit points. Let $x \in E \backslash S$, so that x is not a limit point of S. Therefore, there exists an $r > 0$ such that $S \cap (B(x, r) \backslash \{x\}) = \emptyset$. Hence, $B(x, r) \subset E \backslash S$. This proves that $E \backslash S$ is open, and hence that S is closed. □

Given a set S that is not necessarily closed, we can 'add' points to it until we obtain a closed set. The smallest set that results after we add the minimum amount of points required to make the set closed is called the closure of the set:

Definition 1.5.11. Let E be a Euclidean n-space and let $S \subset E$. The **closure** of S, denoted $\overline{S}$, is the intersection of all closed sets containing S.

It is clear from the above definition that the closure of $S \subset E$ is the smallest closed set containing S, consistent with our above discussion. In fact, the additional points that we need to add to S to make it closed are simply the limit points of S which it does not already contain. Hence, the closure of S is simply the union of S with its set of limit points.

1.6   Completeness

In calculus, we will be working with limits *quite a bit*. In particular, we want to be taking limits of functions. First, as is typical, we will define the limit of a sequence, since limits of functions can be studied by considering limits of sequences:

Definition 1.6.1. Let E be a Euclidean n-space. A sequence $(x_k)_{k=1}^{\infty}$ in E is said to **converge** to $L \in E$ if for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $k \geq N$, $\|x_k - L\| < \epsilon$.

The usual notation to denote the limit of a convergent sequence is $x_k \to L$ as $k \to \infty$, or $\lim_{k \to \infty} x_k = L$. We will use both interchangably.

The metric structure on Euclidean space gives us all of the expected desirable properties about limits and convergence that we know and love from $\mathbb{R}$. If a sequence is convergent, the limit is unique. Closed sets in Euclidean space are exactly the *sequentially closed sets* (i.e. every convergent sequence whose points are all contained within a sequentially closed set have a limit inside the set). Since we have the algebraic structure of a vector space as well, we have some algebraic limit laws: the limit of a sum of sequences is equal to the sum of the limits, and we can pull scalars through limits.

One might notice from the definition of convergence of a sequence that in order to prove directly that a sequence is convergent one must already secretly know what the limit ought to be, so that one can show that the distance between terms in the sequence and the candidate for the limit become arbitrarily small. Guessing what the limit ought to be isn't always a straightforward task, and so fortunately there is another way to show the convergence of a sequence, which relies upon the notion of *completeness*. In a complete metric space, a necessary and sufficient condition for a sequence to be convergent is for the sequence to be *Cauchy*; that is, the terms the sequence progressively become arbitrarily close to *each other*.

Definition 1.6.2. Let E be a Euclidean n-space. A sequence $(x_k)_{k=1}^{\infty}$ in E is said to be **Cauchy** if for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $j, k \geq N$, $\|x_j - x_k\| < \epsilon$.

Intuitively, if the terms in the sequence are gradually varying in distance to a lesser and lesser extent as time goes on, then they must eventually be getting closer to some fixed point, provided they are not clumping around a 'gap'. A complete metric space is one in which there are 'no gaps', much like you would have heard that $\mathbb{R}$ has no gaps.

Definition 1.6.3. Let (X, d) be a metric space. We say that X is **complete** with respect to the metric d if every Cauchy sequence in X converges.

So the big advantage of completeness is that instead of having to find a candidate limit in order to prove a sequence that you suspect converges actually does so, you can just show that the sequence is Cauchy, i.e. the terms just eventually clump up together. Completeness guarantees that the sequence then has a limit. We will now demonstrate the completeness of Euclidean space with respect to the Euclidean metric. The proof is primarily based on the fact that $\mathbb{R}$ is complete, and that Euclidean space is finite-dimensional.

Theorem 1.6.4. *Let E be a Euclidean n-space. Then E is complete with respect to the Euclidean metric d_2.*

Proof. Let $(x_k)_{k=1}^{\infty}$ be a Cauchy sequence in E. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis for E. We can hence express each term of the sequence in terms of the orthonormal basis:

$$x_k = a_{1k}e_1 + \cdots + a_{nk}e_n.$$

We claim that, for each $i = 1, \ldots, n$, the coefficients sequence $(a_{ik})_{k=1}^{\infty}$ is Cauchy in $\mathbb{R}$. To see this, note that

$$
\begin{aligned}
|a_{ij} - a_{ik}| &= \sqrt{(a_{ij} - a_{ik})^2} \\
&\leq \sqrt{\sum_{l=1}^{n}(a_{lj} - a_{lk})^2} \qquad \text{(since the square root function is a real increasing function)} \\
&= \|x_j - x_k\|.
\end{aligned}
$$

Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that for all $j, k \geq N$, $\|x_j - x_k\| < \epsilon$. Such a choice of N is possible since $(x_k)_{k=1}^{\infty}$ is a Cauchy sequence. Then for all $j, k \geq N$, $|a_{ij} - a_{ik}| < \epsilon$. This proves our claim.

Now, by completeness of $\mathbb{R}$, the sequence $(a_{ik})_{k=1}^{\infty}$ converges to, say, $a_i \in \mathbb{R}$. Define $x = a_1e_1 + \cdots + a_ne_n$. We will show that $x_k \to x$ as $k \to \infty$. First, observe that

$$
\begin{aligned}
\|x_k - x\| &= \sqrt{\sum_{i=1}^{n}(a_{ik} - a_i)^2} \\
&\leq \sqrt{n \max_{i=1,\ldots,n}(a_{ik} - a_i)^2} \\
&= \sqrt{n} \max_{i=1,\ldots,n} |a_{ik} - a_i| \\
&< \sqrt{n}\left(\sum_{i=1}^{n}|a_{ik} - a_i|\right).
\end{aligned}
$$

Let $\epsilon > 0$. For each $i = 1, \ldots, n$, choose $N_i \in \mathbb{N}$ such that for all $k \geq N_i$, $|a_{ik} - a_i| < \frac{\epsilon}{n\sqrt{n}}$. Such a choice of N_i is possible since $a_{ik} \to a_i$ as $k \to \infty$. Choose $N = \max\{N_1, \ldots, N_n\}$. Then for all $k \geq N$, we have

$$\|x_k - x\| < \sqrt{n}\sum_{i=1}^{n}\frac{\epsilon}{n\sqrt{n}} = \epsilon.$$

Hence, $x_k \to x$ as $k \to \infty$. This shows that E is complete with respect to the Euclidean metric. $\square$

From the above proof, we can take away a few useful facts. Namely, a sequence in Euclidean space is Cauchy if and only if all of its (real) component sequences are Cauchy, and a sequence in Euclidean space is convergent if and only if all of its (real) component sequences are convergent. Moreover, the components of the limit of the sequence in Euclidean space will simply be the limits of the individual component sequences.

Now we'll consider a few consequences of the completeness of Euclidean space. In fact, most of these properties that we'll look at have corresponding analogues in $\mathbb{R}$. Firstly, we have the Nested Interval Property for Euclidean space, which talks about a nested sequence of closed intervals. By a *closed interval* in

Euclidean space, we mean sets of the form $\{c_1 e_1 + \cdots + c_n e_n \mid c_i \in [a_i, b_i] \,\forall\, i = 1, \ldots, n\}$, where $\{e_1, \ldots, e_n\}$ is an orthonormal basis.

Theorem 1.6.5. *(Nested Interval Property) Let E be a Euclidean n-space. Let $I_1 \supset I_2 \supset \cdots \supset I_k \supset \ldots$ be a decreasing sequence of nonempty closed intervals in E. Then*

$$\bigcap_{k=1}^{\infty} I_k \neq \emptyset.$$

Proof. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis for E, and thus denote the interval I_k by

$$I_k = \{c_1 e_1 + \cdots + c_n e_n \in E \mid c_i \in [a_{ik}, b_{ik}] \,\forall\, i = 1, \ldots n\}.$$

Since $I_1 \supset I_2 \supset \ldots I_k \supset \ldots$ is a decreasing sequence of closed intervals, we have that for all $i = 1, \ldots, n$, $(a_{ik})_{k=1}^{\infty}$ is an increasing sequence of real numbers. Furthermore, the sequence is bounded above by $b_{i1} \in \mathbb{R}$. Hence, we have a bounded, monotonic sequence of real numbers. By the Monotone Convergence Theorem in $\mathbb{R}$, the sequence converges: so $a_{ik} \to a_i$ as $k \to \infty$. We will hence show that $x = a_1 e_1 + \cdots + a_n e_n \in \bigcap_{k=1}^{\infty} I_k$.

Let $m \in \mathbb{N}$. Fix $i \in \{1, \ldots, n\}$. Since $(a_{ik})_{k=1}^{\infty}$ is an increasing sequence, then $a_{im} \leq a_i$. Additionally, since $a_{ik} \leq b_{im}$ for all $k \in \mathbb{N}$, then $a_i \leq b_{im}$, since limits preserve inequalities. Hence, we have that $a_i \in [a_{im}, b_{im}]$. Since this holds for each i, this shows that $x \in I_m$. Since the choice of $m \in \mathbb{N}$ is arbitrary, then x is in every closed interval I_m, and thus in the intersection of all of them. $\qquad \square$

The Nested Interval Property is an immediate consequence of the completeness of Euclidean space. However, one may glance at our above proof and wonder where we used the completeness of Euclidean space in the argument? Secretly, we kind of cheated: rather than appealing to completeness of Euclidean space directly in the proof, we made use of the completeness of $\mathbb{R}$ (in the form of the Monotone Convergence Theorem) and the finite-dimensionality of Euclidean space, which were the two key facts exploited in the completeness proof. One could certainly prove the Nested Interval Property directly from the completeness of Euclidean space, but we thought the above proof was quite simple and a further demonstration of how easy things are in a finite dimensional setting.

Next up, we'll look at another key consequence of completeness: the Bolzano-Weierstrass Theorem. This theorem states that every bounded sequence has a convergent subsequence. First, let's define what it means for a set to be bounded in Euclidean space:

Definition 1.6.6. Let E be a Euclidean n-space. A subset $A \subset E$ is **bounded** if there exists an $M > 0$ such that for all $x \in A$, $\|x\| \leq M$.

A bounded sequence in E would thus be a sequence $(x_k)_{k=1}^{\infty}$ in which the set of its terms $\{x_k \in E \mid k \in \mathbb{N}\}$ is bounded. First, let's prove a slightly rephrased version of the Bolzano-Weierstrass Theorem, which makes use of the Nested Interval Property.

Proposition 1.6.7. *Every bounded infinite set in Euclidean n-space has at least one limit point.*

Proof. Let E be a Euclidean n-space and let $A \subset E$ be a bounded infinite set. Then there exists an $M > 0$ such that for all $x \in A$, $\|x\| \leq M$. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis in E. Define the closed interval $I_1 = \{a_1 e_1 + \cdots + a_n e_n \in E \mid a_i \in [-M, M] \,\forall\, i = 1, \ldots n\}$. Clearly, $A \subset I_1$. Then partition I_1 into 2^n closed intervals of side length $\frac{M}{2}$. Choose I_2 to be one of the subintervals that contains infinitely many points of A.

Continue in this fashion: given the closed interval I_k, partition it into 2^n subintervals of side length $\frac{M}{2^k}$ and choose I_{k+1} to be a subinterval which contains infinitely many points of A. This process constructs a decreasing sequence $I_1 \supset I_2 \supset \cdots \supset I_k \supset \ldots$ of nonempty closed intervals. Hence, by the Nested Interval Property, $\bigcap_{k=1}^{\infty} I_k \neq \emptyset$. Let $x \in \bigcap_{k=1}^{\infty} I_k$. We will show that x is a limit point of A.

Let $\epsilon > 0$. Choose $k \in \mathbb{N}$ such that $\frac{1}{k} < \frac{\epsilon}{M\sqrt{n}}$. Such a choice of k is possible since $\mathbb{R}$ satisfies the Archimedean property. Choose $a \in A$ such that $a \in I_k$, which is possible since I_k contains infinitely many points of A. Since we also have that $x \in I_k$, then the (Euclidean) distance between x and a must be less

than or equal to the largest possible distance between two points in the closed interval. For a closed cubical interval of equal side lengths $r > 0$, the greatest distance is $\sqrt{nr^2} = r\sqrt{n}$. Hence, we have that

$$\|x - a\| \leq \frac{M}{2^{k-1}}\sqrt{n} \leq \frac{M}{k}\sqrt{n} < \epsilon.$$

Hence, from the above expression we have that $a \in A \cap (B(x, \epsilon)\backslash\{x\})$. Since this is true for all $\epsilon > 0$, it follows that x is a limit point of A. $\square$

Now the *actual* Bolzano-Weierstrass Theorem follows as an easy consequence of the above result:

Theorem 1.6.8. *(Bolzano-Weierstrass Theorem) Every bounded sequence in Euclidean n-space has a convergent subsequence.*

Proof. Let E be a Euclidean n-space and let $(x_k)_{k=1}^{\infty}$ be a bounded sequence in E. Then there exists an $M > 0$ such that $\|x_k\| \leq M$ for all $k \in \mathbb{N}$. Thus, the set $\{x_k \in E \mid k \in \mathbb{N}\}$ is a bounded subset of E. Thus, by Proposition 1.6.7, it has a limit point, say $x \in E$.

Since x is a limit point of the set $\{x_k \in E \mid k \in \mathbb{N}\}$, then for every $m \in \mathbb{N}$, pick an $x_{k_m} \in \{x_k \in E \mid k \in \mathbb{N}\} \cap \left(B\left(x, \frac{1}{m}\right)\backslash\{x\}\right)$ such that $k_m > k_{m'}$ for all $m' < m$. This constructs the subsequence $(x_{k_m})_{m=1}^{\infty}$, which converges to x as we will now show.

Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that $\frac{1}{N} < \epsilon$. Such a choice of N is possible since $\mathbb{R}$ satisfies the Archimedean property. Then for all $m \geq N$, we have

$$\|x_{k_m} - x\| < \frac{1}{m} < \frac{1}{N} < \epsilon.$$

Thus, $x_{k_m} \to x$ as $m \to \infty$. $\square$

1.7   Continuity

We will now finally define the notion of limits of functions between Euclidean spaces, which will also give us the necessary framework to talk about continuous functions.

Definition 1.7.1. Let E, F be Euclidean spaces, $U \subset E$ a subset of E, $x_0 \in E$ a limit point of U and $f : U \to F$ a function. The limit of f as x approaches x_0 is the point $y_0 \in F$ such that for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in U\backslash\{x_o\}$ with $\|x - x_0\| < \delta$ we have $\|f(x) - y_0\| < \epsilon$, if such a point y_0 exists. Else, we say that the limit does not exist.

The notations $f(x) \to y_0$ as $x \to x_0$ and $\lim_{x \to x_0} f(x) = y_0$ are common - we will employ both throughout this text.

The astute reader might notice that we are being somewhat sloppy, as we are using $\|-\|$ to denote both the Euclidean norms in E and F. However, it should be fairly clear from the context which norm is being used.

A very useful fact is that limits of functions can be characterised in terms of limits of sequences, the latter of which are usually much simpler to deal with (we can use all of our facts about limits of sequences and apply them to limits of functions).

Proposition 1.7.2. *Let E, F be Euclidean spaces, $U \subset E$ a subset of E, $x_0 \in E$ a limit point of U and $f : U \to F$ a function. Then $\lim_{x \to x_0} f(x) = y_0$ if and only if for every sequence $(x_k)_{k=1}^{\infty}$ in $U\backslash\{x_0\}$ with $x_k \to x_0$ as $k \to \infty$, we have that $f(x_k) \to y_0$ as $k \to \infty$.*

Proof. ($\implies$) Suppose $\lim_{x \to x_0} f(x) = y_0$. Let $(x_k)_{k=1}^{\infty}$ be a sequence in $U\backslash\{x_0\}$ with $x_k \to x_0$ as $k \to \infty$. Consider the sequence $(f(x_k))_{k=1}^{\infty}$ in F. We will show that $f(x_k) \to y_0$ as $k \to \infty$.

Let $\epsilon > 0$. Choose $\delta > 0$ such that for all $x \in U\backslash\{x_0\}$ with $\|x - x_0\| < \delta$ we have $\|f(x) - y_0\| < \epsilon$. Such a choice of δ is possible since $\lim_{x \to x_0} f(x) = y_0$. Choose $N \in \mathbb{N}$ such that for all $k \geq N$, $\|x_k - x_0\| < \delta$. Such

a choice of N is possible since $(x_k)_{k=1}^\infty$ converges to x_0. Then for all $k \geq N$, we have that $\|f(x_k) - y_0\| < \epsilon$. This proves that $f(x_k) \to y_0$ as $k \to \infty$.

($\impliedby$) Suppose that for every sequence $(x_k)_{k=1}^\infty$ in $U \backslash \{x_0\}$ with $x_k \to x_0$ as $k \to \infty$, we have that $f(x_k) \to y_0$ as $k \to \infty$. Assume for a contradiction that $\lim_{x \to x_0} f(x) \neq y_0$. Then there exists an $\epsilon > 0$ such that for all $k \in \mathbb{N}$, there exists an $x_k \in U \backslash \{x_0\}$ with $\|x_k - x_0\| < \frac{1}{k}$ but $\|f(x_k) - y_0\| \geq \epsilon$. This constructs a sequence $(x_k)_{k=1}^\infty$ in $U \backslash \{x_0\}$.

Clearly, $x_k \to x_0$ as $k \to \infty$. Hence, $f(x_k) \to y_0$ as $k \to \infty$. This means that there exists an $N \in \mathbb{N}$ such that for all $k \geq N$, we have that $\epsilon \leq \|f(x_k) - y_0\| < \epsilon$. This is impossible, so we are forced to conclude that $\lim_{x \to x_0} f(x) = y_0$. $\qquad \square$

We can thus easily derive the algebraic limit laws of functions from the above result and the corresponding facts for sequences: the limit of a sum of functions is the sum of the limits, and you can pull scalars through limits. Moreover, limits behave 'nicely' with composition of functions: given $f : E \to F$ and $g : F \to G$ and $a \in E$, if $f(x) \to L \in F$ as $x \to a$ and $g(y) \to M$ as $y \to L$, then $g \circ f(x) \to M$ as $x \to a$.

We will now introduce the notion of a continuous function between Euclidean spaces. Intuitively, the idea is that sufficiently small variations in the domain result in arbitrarily small variations in the image of the function.

Definition 1.7.3. Let E, F be Euclidean spaces, $U \subset E$ a subset of E and $f : U \to F$ a function. We say that f is **continuous** at $x_0 \in E$ if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in U$ with $\|x - x_0\| < \delta$, we have $\|f(x) - f(x_0)\| < \epsilon$.

If f is continuous at all $x \in U$, then we say that f is continuous on U.

One should notice that the above definition closely resembles that of the limit of a function, instead that the limit of the function as you approach a point is equal to the function evaluated at that particular point, if it is continuous there. This is why people often like to say that a function $f : U \to F$ is continuous at $x_0 \in U$ if $\lim_{x \to x_0} f(x) = f(x_0)$. As continuity can be formulated in terms of a limit, we can obtain an analogous result to Proposition 1.7.2, where we can formulate continuity via limits of sequences. The proof follows the same argument as Proposition 1.7.2.

Proposition 1.7.4. *Let E, F be Euclidean spaces, $U \subset E$ a subset of E and $f : U \to F$ a function. Then f is continuous at $x_0 \in U$ if and only if for all sequences $(x_k)_{k=1}^\infty$ in U with $x_k \to x_0$ as $k \to \infty$, we have that $f(x_k) \to f(x_0)$ as $k \to \infty$.*

Thus, by the algebraic limit laws, the sum of continuous functions is continuous, and scalar multiples of continuous functions are continuous. Additionally, the composition of continuous functions is continuous.

Let E be a Euclidean n-space and F a Euclidean m-space. Let $\bar{e}_1, \ldots, \bar{e}_m$ be a basis for F. A multivariable function $f : E \to F$ can be studied in terms of its real-valued component functions with respect to the basis in F: $f(x) = f_1(x)\bar{e}_1 + \cdots + f_m(x)\bar{e}_m$. One nice fact is that f is continuous if and only if each of the component functions $f_i : E \to \mathbb{R}$, $i = 1, \ldots, m$, are continuous, so that we can simply deal with real-valued functions.

Proposition 1.7.5. *A function $f : E \to F$ is continuous at $x_0 \in E$ if and only if all of its component functions $f_i : E \to \mathbb{R}$ are continuous at x_0.*

Proof. ($\implies$) Suppose f is continuous at x_0. Let $i \in \{1, \ldots, m\}$ and consider the component map $f_i : E \to \mathbb{R}$. Then for all $x \in E$, we have

$$|f_i(x) - f_i(x_0)| \leq \left(\sum_{k=1}^m (f_k(x) - f_k(x_0))^2 \right)^{\frac{1}{2}} = \|f(x) - f(x_0)\|.$$

Taking the limit as $x \to x_0$ yields

$$0 \leq \lim_{x \to x_0} |f_i(x) - f_i(x_0)| \leq 0$$

since $\lim_{x \to x_0} \| f(x) - f(x_0) \|$ by continuity of f at x_0. It follows that f_i is continuous at x_0.

($\Longleftarrow$) Suppose $f_i : E \to \mathbb{R}$ is continuous for each $i \in \{1, \dots m\}$. Then for all $x \in E$, we have

$$\| f(x) - f(x_0) \|^2 = \sum_{k=1}^{m} (f_k(x) - f_k(x_0))^2$$
$$\leq m \times \max_{i \in \{1, \dots, m\}} (f_i(x) - f_i(x_0))^2$$

This, in turn, gives

$$\| f(x) - f(x_0) \| \leq \sqrt{m} \times \max_{i \in \{1, \dots, m\}} |f_i(x) - f_i(x_0)| \leq \sqrt{m} \sum_{k=1}^{m} |f_k(x) - f_k(x_0)|.$$

Applying the limit as $x \to x_0$ yields

$$0 \leq \lim_{x \to x_0} \| f(x) - f(x_0) \| \leq \sqrt{m} \sum_{k=1}^{m} \lim_{x \to x_0} |f_k(x) - f_k(x_0)| = 0$$

which follows from the continuity of each of the f_k at x_0. Hence, we have that $\lim_{x \to x_0} \| f(x) - f(x_0) \| = 0$, which implies that f is continuous at x_0. $\qquad\square$

We conclude this section with a fun fact: linear maps between Euclidean spaces are automatically continuous. So linear maps are really, really nice to work with! This is another consequence of the finite dimensionality of the spaces we are working in.

Theorem 1.7.6. *Let $f : E \to F$ be a linear mapping between Euclidean spaces. Then f is continuous.*

Proof. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis for E. Let $x_0 \in E$, and let $(x_k)_{k=1}^{\infty}$ be a sequence in E such that $x_k \to x_0$ as $k \to \infty$. We can write each term of the sequence as a linear combination of the basis vectors

$$x_k = a_{1k} e_1 + \cdots + a_{nk} e_n.$$

Let's also expand x_0 in terms of the basis vectors

$$x_0 = a_1 e_1 + \cdots + a_n e_n.$$

Now, we know that $x_k \to x_0$ as $k \to \infty$ if and only if all of the real component sequences $(a_{ik})_{k=1}^{\infty}$ converge to the corresponding component a_i of x_0, i.e. we have that $a_{ik} \to a_i$ as $k \to \infty$ for each $i = 1, \dots, n$.

Finally, by the linearity of f, we have that

$$\lim_{k \to \infty} f(x_k) = \lim_{k \to \infty} \sum_{i=1}^{n} a_{ik} f(e_i)$$
$$= \sum_{i=1}^{n} f(e_i) \lim_{k \to \infty} a_{ik} \qquad \text{(by linearity of limits)}$$
$$= \sum_{i=1}^{n} a_i f(e_i)$$
$$= f\left(\sum_{i=1}^{n} a_i e_i \right) \qquad \text{(by linearity of } f\text{)}$$
$$= f(x_0).$$

By Proposition 1.7.4, this shows that f is continuous at x_0. Since the choice of $x_0 \in E$ was arbitrary, it follows that f is continuous on all of E. $\qquad\square$

1.8   Compactness

Here we introduce a property of subsets of Euclidean space called *compactness*. In some sense, it is an extension of the notion of finiteness of sets. While a compact set does not necessarily have finitely many elements, in a way it can always be described with a 'finite amount of information'. The most general setting in which one can define compactness is in a topological space, in which the notion of a set being compact is basically that the set is 'topologically finite':

Definition 1.8.1. Let X be a topological space. A subset $A \subset X$ is **compact** if for every collection of open sets $\mathcal{U} = \{U_\alpha \mid \alpha \in I\}$ satisfying $A \subset \bigcup_{\alpha \in I} U_\alpha$, there exists a finite subset $\{U_1, \dots, U_r\} \subset \mathcal{U}$ such that $A \subset U_1 \cup \cdots \cup U_r$.

A collection of open sets $\mathcal{U}$ in the definition above is called an **open cover** of the set A. Thus, we can more concisely recite the above definition by stating that a set is compact if every open cover has a finite subcover. So if we try to write down a compact set as a union of infinitely many open sets, we will always have 'redundant information', and so we can afford to throw away most of the sets from the union until finitely many remain.

Any finite subset of a topological space is compact. This is in alignment with our introductory spiel that compactness generalises the notion of finiteness for sets.

Since Euclidean space is also a metric space, there exists a different notion of compactness that is phrased in terms of convergence of sequences, called *sequential compactness*.

Definition 1.8.2. Let X be a metric space. A subset $A \subset X$ is **sequentially compact** if every sequence in A has a convergent subsequence whose limit is contained in A.

We also have that every finite subset of a metric space is sequentially compact. Any sequence whose terms are taken from a finite set must take on at least one of the values from the set infinitely many times, and so we can construct from it a constant subsequence, which is obviously convergent within the set. So sequential compactness is another possibility for extending the notion of finiteness in a metric space setting.

Since every metric space can also be considered as a topological space using the metric topology, one can define both the notions of compactness and sequential compactness on a metric space. The remarkable fact that we will proceed to show is that these notions actually coincide.

Before we tackle the proof, we will require a useful auxillary result that we will use in order to construct a finite subcover of a given open cover of a sequentially compact set.

Lemma 1.8.3. *(Lebesgue's number lemma) Let $A \subset X$ be a sequentially compact subset of a metric space X and let $\mathcal{U}$ be an open cover of A. Then there exists a $\delta > 0$ such that for all $x \in A$, there exists a $U \in \mathcal{U}$ such that $B(x, \delta) \subset U$.*

Proof. Assume for a contradiction that for all $\delta > 0$, there exists an $x \in A$ such that $B(x, \delta) \cap (X \backslash U) \neq \emptyset$ for all $U \in \mathcal{U}$. Then for each $n \in \mathbb{N}$, pick an $x_n \in A$ such that $B\left(x_n, \frac{1}{n}\right) \cap (X \backslash U) \neq \emptyset$ for all $U \in \mathcal{U}$.

This defines a sequence $(x_n)_{n=1}^\infty$ in A, which has a convergent subsequence $(x_{n_k})_{k=1}^\infty$ since A is sequentially compact. Suppose that $x_{n_k} \to a$ as $k \to \infty$. Since $a \in A$, then $a \in V$ for some open set $V \in \mathcal{U}$, since $\mathcal{U}$ is an open cover of A. This means that there exists an $\epsilon > 0$ such that $B(a, \epsilon) \subset V$. But by convergence of the subsequence $(x_{n_k})_{k=1}^\infty$, there exists an $N_1 \in \mathbb{N}$ such that for all $k \geq N_1$ $d(x_{n_k}, a) < \frac{\epsilon}{2}$. Additionally, by the Archimedean property, there exists an $N_2 \in \mathbb{N}$ such that for all $k \geq N_2$, $\frac{1}{k} < \frac{\epsilon}{2}$. Choose $N = \max\{N_1, N_2\}$. Then for all $k \geq N$, we have that for any $y \in B\left(x_{n_k}, \frac{1}{k}\right)$:

$$d(y, a) \leq d(y, x_{n_k}) + d(x_{n_k}, a) < \frac{1}{k} + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

So $y \in B(a, \epsilon)$, and hence $B\left(x_{n_k}, \frac{1}{k}\right) \subset B(a, \epsilon) \subset V$, which contradicts the choice of x_{n_k}. $\square$

The positive constant $\delta > 0$ in the lemma above is called a *Lebesgue number* for the open cover $\mathcal{U}$.

Theorem 1.8.4. *Let X be a metric space and let $A \subset X$. Then A is compact if and only if it is sequentially compact.*

Proof. ($\implies$) We prove the contrapositive. Let $(x_n)_{n=1}^\infty$ be a sequence in A with no convergent subsequence, and denote by $S = \{x_n \in A \mid n \in \mathbb{N}\}$ to be the set of terms in the sequence. Then any $a \in A$ cannot be a limit point of S, else one could construct a subsequence of $(x_n)_{n=1}^\infty$ that converges to a. Thus, for all $a \in A$, there exists an $r_a > 0$ such that $(B(a, r_a)\backslash\{a\}) \cap S = \emptyset$, that is, the ball $B(a, r_a)$ contains no points in the sequence $(x_n)_{n=1}^\infty$, except possibly a itself. Clearly, we have that $\mathcal{U} = \bigcup_{a \in A} B(a, r_a)$ is an open cover of A. The open cover $\mathcal{U}$ cannot have a finite subcover, since any finite subset of $\mathcal{U}$ can only contain finitely many terms of the sequence $(x_n)_{n=1}^\infty$, all of which are points belonging to A. We have thus proven that if A is *not* sequentially compact, then it is *not* compact.

($\impliedby$) Suppose A is sequentially compact. Let $\mathcal{U}$ be an open cover of A. By Lebesgue's number lemma, there exists a Lebesgue number $\delta > 0$ such that for all $x \in A$, there exists an open set $U \in \mathcal{U}$ such that $B(x, \delta) \subset U$. We will use this property to construct a finite subcover of $\mathcal{U}$.

Pick any $U_1 \in \mathcal{U}$. If $A \subset U_1$, then we are done (U_1 would be a finite subcover of $\mathcal{U}$). Else, pick any $x_1 \in A\backslash U_1$. By Lebesgue's number lemma, we can pick $U_2 \in \mathcal{U}$ such that $B(x_1, \delta) \subset U_2$. If $A \subset U_1 \cup U_2$, we are done. Else, pick an $x_2 \in A\backslash(U_1 \cup U_2)$ and $U_3 \in \mathcal{U}$ satisfying $B(x_2, \delta) \subset U_3$. Notice that since $x_2 \notin U_1 \cup U_2$, then $x_2 \notin B(x_1, \delta)$, so that $d(x_1, x_2) \geq \delta$.

Continue the process in the following fashion. Suppose we had a finite collection of open sets $U_1, \ldots U_n \in \mathcal{U}$ that was *not* a subcover of $\mathcal{U}$ and a finite subset $\{x_1, \ldots, x_{n-1}\} \subset A$ such that $d(x_i, x_j) \geq \delta$ for each $i, j = 1, \ldots, n-1$ with $i \neq j$. Pick $x_n \in A\backslash(U_1 \cup \cdots \cup U_n)$ and $U_{n+1} \in \mathcal{U}$ satisfying $B(x_n, \delta) \subset U_{n+1}$. Since $x_n \notin (U_1 \cup \cdots \cup U_n)$, then $x_n \notin U_i$ for each $i = 1, \ldots, n-1$, and hence $x_n \notin B(x_i, \delta)$ so that $d(x_n, x_i) \geq \delta$.

Suppose that this process did *not* terminate. That is, there does not exist an $n \in \mathbb{N}$ such that $A \subset U_1 \cup \cdots \cup U_n$. Then by induction on $\mathbb{N}$, we obtain a sequence $(x_n)_{n=1}^\infty$ in A with $d(x_i, x_j) \geq \delta$ for all $i, j \in \mathbb{N}$ with $i \neq j$. By sequential compactness of A, this sequence has a convergent subsequence $(x_{n_k})_{k=1}^\infty$ with limit $x \in A$. Hence, there exists an $N \in \mathbb{N}$ such that for all $k \geq N$, $d(x_{n_k}, x) < \frac{\delta}{2}$. But by construction, we have that $x_{n_l} \notin B(x_k, \delta)$ for all $l \geq k$, so $d(x_{n_k}, x_{n_l}) \geq \delta$. Taking $l \to \infty$, by continuity of the metric we have that $d(x_{n_k}, x) \geq \delta$. This is a contradiction, and so we must have that there exists an $n \in \mathbb{N}$ such that $A \subset U_1 \cup \cdots \cup U_n$, so that $\mathcal{U}$ has a finite subcover. $\qquad\square$

In light of the above result, we will simply use the term 'compactness' to refer to both topological compactness and sequential compactness in the context of subsets of Euclidean space. In practice, most of the time when we invoke the property of compactness for a subset of Euclidean space, we will use sequential compactness because sequences are easier to work with, owing to the rich theory of sequences and convergence in a metric space.

Even better, the conditions for compactness become significantly simpler in a finite dimensional setting. This is the celebrated *Heine-Borel Theorem*, which states that the compact subsets of Euclidean space are exactly the closed and bounded subsets. In general, for a metric space, every compact set is closed and bounded, but in Euclidean space, this becomes a sufficient condition for compactness.

Theorem 1.8.5. *(Heine-Borel Theorem) Let E be a Euclidean space and let $A \subset E$. Then A is compact if and only if it is closed and bounded.*

Proof. ($\implies$) Suppose $A \subset E$ is compact. To see that A is closed, let $(a_n)_{n=1}^\infty$ be a sequence in A with limit $a \in E$. Since A is (sequentially) compact, there exists a convergent subsequence $(a_{n_k})_{k=1}^\infty$ with limit $x \in A$. However, since the original sequence $(a_n)_{n=1}^\infty$ is convergent, then the limits of any of its subsequences must be equal to a. Hence, we see that $a = x$, so that $a \in A$. This shows that A is *sequentially closed*, and hence closed.

Now, suppose for a contradiction that A was unbounded. Then for each $n \in \mathbb{N}$, we can choose an $x_n \in A$ such that $\|x_n\| > n$. This constructs a sequence $(x_n)_{n=1}^\infty$ in A. By (sequential) compactness of A, this sequence has a convergent subsequence $(x_{n_k})_{k=1}^\infty$ with limit $x \in A$. However, taking $N = \lceil \|x\| + 1 \rceil$, we have

for all $k \geq N$ that $\|x_{n_k} - x\| \geq \|\|x_{n_k}\| - \|x\|\| \geq 1$ by the reverse triangle inequality, which contradicts the fact that $x_{n_k} \to x$ as $k \to \infty$. Hence, we must have that A is bounded.

($\Longleftarrow$) Suppose $A \subset E$ is closed and bounded. Let $(a_n)_{n=1}^\infty$ be a sequence in A. Thus, $(a_n)_{n=1}^\infty$ is a bounded sequence. By the Bolzano-Weierstrass Theorem, there exists a convergent subsequence $(a_{n_k})_{k=1}^\infty$ with limit $a \in E$. However, since A is closed, we must have that the limit is contained in A, i.e. $a \in A$. This proves the (sequential) compactness of A. $\square$

Now that we have seen the various ways we can characterise a compact set in Euclidean space, we will now spend some time looking at why compact sets are nice to work with. In particular, compact sets interact nicely with continuous functions. The first fact we will look at is that if a continuous real-valued function defined on a compact set is always positive, then it does not get arbitrarily close to 0, that is, the function is bounded below by a positive constant.

Proposition 1.8.6. *Let E be a Euclidean space, $A \subset E$ a compact set, and $f : A \to \mathbb{R}$ a continuous function with $f(x) > 0$ for all $x \in A$. Then there exists an $\alpha > 0$ such that $f(x) > \alpha$ for all $x \in A$.*

Proof. Suppose for a contradiction that for all $\epsilon > 0$, there exists an $x \in A$ such that $f(x) \leq \epsilon$. Then for each $n \in \mathbb{N}$, choose an $x_n \in A$ such that $f(x_n) \leq \frac{1}{n}$. This defines a sequence $(x_n)_{n=1}^\infty$ in A. By compactness of A, there exists a subsequence $(x_{n_k})_{k=1}^\infty$ that converges to $a \in A$. By continuity of f, we thus have that $f(x_{n_k}) \to f(a)$ as $k \to \infty$. But

$$\lim_{k \to \infty} f(x_{n_k}) \leq \lim_{k \to \infty} \frac{1}{n_k}$$
$$\therefore f(a) \leq 0.$$

This contradicts the fact that $f(x) > 0$ for all $x \in A$. $\square$

We will make use of the above result in the following section.

Compact sets are preserved by continuous functions. That is, the image of a compact set under a continuous function is itself a compact set:

Proposition 1.8.7. *Let $f : E \to F$ be a continuous mapping between Euclidean spaces and let $A \subset E$ be a compact set. Then $f(A)$ is a compact subset of F.*

Proof. Let $(y_n)_{n=1}^\infty$ be a sequence in $f(A)$. Then for each $n \in \mathbb{N}$, there exists an $x_n \in A$ such that $y_n = f(x_n)$. Consider the sequence $(x_n)_{n=1}^\infty$ in A. By compactness of A, this sequence has a convergent subsequence $(x_{n_k})_{k=1}^\infty$ with limit $x \in A$. By continuity of f, we have that $f(x_{n_k}) \to f(x)$ as $k \to \infty$, i.e. $y_{n_k} \to f(x) \in f(A)$ as $k \to \infty$. So $(y_{n_k})_{k=1}^\infty$ is a convergent subsequence of $(y_n)_{n=1}^\infty$ whose limit is contained in $f(A)$. Hence, $f(A)$ is compact. $\square$

An immediate application of the above fact is the *Extreme Value Theorem*. This is an existence result that guarantees that a continuous real-valued function defined on a compact set achieves global extrema on its domain. Since one of the main goals of differential calculus is optimisation, it is a good first step to know that global extrema exist somewhere so it makes sense to look for them in the first place!

Theorem 1.8.8. *(Extreme Value Theorem) Let E be a Euclidean space, $A \subset E$ a compact set and $f : A \to \mathbb{R}$ a continuous function. Then f attains a maximum and a minimum on A.*

Proof. By Proposition 1.8.7, we have that $f(A)$ is a compact set. By the Heine-Borel Theorem, this implies that $f(A)$ is a closed and bounded subset of $\mathbb{R}$. The boundedness of $f(A)$ implies that $\sup f(A) < \infty$ and $\inf f(A) > -\infty$. For each $n \in \mathbb{N}$, choose $x_n, y_n \in A$ such that $\sup f(A) - \frac{1}{n} < f(x_n) \leq \sup f(A)$ and $\inf f(A) \leq f(y_n) < \inf f(A) + \frac{1}{n}$. By taking limits as $n \to \infty$ and applying the Squeeze Theorem, we see that $f(x_n) \to \sup f(A)$ and $f(y_n) \to \inf f(A)$ as $n \to \infty$. By closedness of $f(A)$, we have that $\sup f(A) \in f(A)$ and $\inf f(A) \in f(A)$, so there exists $x_0, x_1 \in A$ satisfying $f(x_0) = \sup f(A)$ and $f(x_1) = \inf f(A)$, i.e. f attains a maximum at x_0 and a minimum at x_1. $\square$

1.9   Equivalence of Norms

In Section 1.2, we saw that the Euclidean inner product *uniquely* induces the Euclidean norm $\|-\|_2$ on $\mathbb{R}^n$, which gave us sufficient machinery to define distances, convergence, continuity, etc. If we (temporarily) forget about the inner product structure of Euclidean space, a very natural thing to think about is what would happen if we equipped $\mathbb{R}^n$ with a *different* norm. Would the resulting topological structure differ from that of Euclidean space? Perhaps surprisingly, the answer is no. While different norms may prescribe different lengths to the same vector, they actually all induce the exact same topology in a finite dimensional setting!

To discuss this idea further on a more solid ground, we will define what it means in general for two norms to be 'the same' under a topological lens. We will call topologically compatible pairs of norms to be *strongly equivalent*:

Definition 1.9.1. Let V be a vector space and let $\|-\|_1$ and $\|-\|_2$ be norms on V. We say that the norms $\|-\|_1$ and $\|-\|_2$ are **strongly equivalent** if there exists constants $C, D > 0$ such that $C\|x\|_1 \leq \|x\|_2 \leq D\|x\|_1$ for all $x \in V$.

Strong equivalence of norms on a vector space defines an equivalence relation on all of the possible norms that can be defined on the vector space.

The condition for two norms being strongly equivalent to one another may seem somewhat opaque at a first glance. Admittedly, it is more of an *operational definition* in that the definition itself is not very intuitive or particularly illuminating, but it makes proving the shared properties of strongly equivalent norms a relatively simple task thanks to the nice inequality involving the two different norms that it establishes. We will say more on the implications of having strongly equivalent norms a bit later in this section. First, we will establish the very important fact that any norm that one can define on $\mathbb{R}^n$ is strongly equivalent to any other.

Theorem 1.9.2. *Let E be a Euclidean n-space. Then any two norms on E are strongly equivalent.*

Proof. We will show that every norm on E is strongly equivalent to the Euclidean norm $\|-\|$. Let $\|-\|_0$ be any norm on E. Define the function $f : E \backslash \{0\} \to \mathbb{R}$ by

$$f(x) = \frac{\|x\|}{\|x\|_0}.$$

Note that $f(x) > 0$ for all $x \in E \backslash \{0\}$.

Consider the unit sphere in the Euclidean norm, i.e. the set $S = \{x \in E \mid \|x\| = 1\}$. The set $S \subset E$ is compact since it is a closed, bounded set (Heine-Borel Theorem). Hence, if we restrict the domain of f to S, we can apply Proposition 1.8.6 and obtain a constant $C > 0$ such that $f(x) > C$ for all $x \in S$. Hence, we have that $\frac{\|x\|}{\|x\|_0} > C$ for all $x \in E$ with $\|x\| = 1$.

Now let $y \in E$ with $y \neq 0$. Define $\hat{\mathbf{y}} = \frac{y}{\|y\|}$; the vector y normalised to unit length in the Euclidean norm. Clearly, $\|\hat{\mathbf{y}}\| = 1$, so that we have

$$\frac{\|\hat{\mathbf{y}}\|}{\|\hat{\mathbf{y}}\|_0} > C.$$

But by homogeneity of the norm, we can pull a factor of $\frac{1}{\|y\|}$ outside the norms in both the numerator and denominator of the left-hand-side of the above expression, which cancel. We hence arrive at

$$\frac{\|y\|}{\|y\|_0} > C.$$

This rearranges to $C\|y\|_0 < \|y\|$, which holds for all $y \neq 0$. If $y = 0$, then we have a trivial equality $(0 = C\|y\|_0 = \|y\| = 0)$. Hence, we have that $C\|x\|_0 \leq \|x\|$ for all $x \in E$.

Similarly, define the function $g : E \backslash \{0\} \to \mathbb{R}$ by

$$g(x) = \frac{\|x\|_0}{\|x\|}.$$

We have that $g(x) > 0$ for all $x \in E \backslash \{0\}$. Once again, we restrict the domain of g to the unit sphere S and apply Proposition 1.8.6 to obtain a constant $D' > 0$ such that $g(x) > D'$ for all $x \in S$. Hence, we have that $\frac{\|x\|_0}{\|x\|} > D'$ for all $x \in E$ with $\|x\| = 1$.

Let $y \in E$ with $y \neq 0$. Then

$$\frac{\|\hat{\boldsymbol{y}}\|_0}{\|\hat{\boldsymbol{y}}\|} > D'$$

$$\frac{\left\| \frac{y}{\|y\|} \right\|_0}{\left\| \frac{y}{\|y\|} \right\|} > D'$$

$$\frac{\|y\|}{\|y\|} \frac{\|y\|_0}{\|y\|} > D'$$

$$\frac{\|y\|_0}{\|y\|} > D'.$$

Hence, we have that $\|y\| < D\|y\|_0$ for all $y \neq 0$, where we have defined $D = \frac{1}{D'} > 0$. Once again, if $y = 0$, then we trivially have that $\|y\| = D\|y\|_0$. Hence, for any $x \in E$, we have the inequality $\|x\| \leq D\|x\|_0$.

Putting this all together, we have that $C\|x\|_0 \leq \|x\| \leq D\|x\|_0$ for all $x \in E$. This proves that the norm $\|-\|_0$ is strongly equivalent to the Euclidean norm. Since strong equivalence is an equivalence relation on the norms on E, it follows that every norm on E is strongly equivalent to one another. $\qquad\square$

We are now ready to explore the properties shared by strongly equivalent norms. First, we formally prove the statement that strongly equivalent norms induce the same topology by showing that the open sets in the metric topology induced by $\|-\|_1$ exactly coincide with the open sets in the metric topology induced by $\|-\|_2$.

Proposition 1.9.3. *Let V be a vector space and let $\|-\|_1$ and $\|-\|_2$ be strongly equivalent norms on V. Then a subset $U \subset V$ is open with respect to the norm $\|-\|_1$ if and only if U is open with respect to the norm $\|-\|_2$.*

Proof. ($\Longrightarrow$) Suppose U is open with respect to $\|-\|_1$, i.e. for all $x \in U$, there exists an $r_x > 0$ such that $B_1(x, r_x) = \{y \in V \mid \|x - y\|_1 < r_x\} \subset U$.

Since $\|-\|_1$ and $\|-\|_2$ are strongly equivalent, there exists $C > 0$ such that $C\|v\|_1 \leq \|v\|_2$ for all $v \in V$. Let $x \in U$ and let $y \in B_2(x, Cr_x) = \{z \in V \mid \|x - z\|_2 < Cr_x\}$. Then

$$C\|x - y\|_1 \leq \|x - y\|_2 < Cr_x.$$

We hence obtain $\|x - y\|_1 < r_x$, which demonstrates that $y \in B_1(x, r_x)$. We thus have that $B_2(x, Cr_x) \subset B_1(x, r_x) \subset U$, which proves that U is open with respect to $\|-\|_2$.

($\Longleftarrow$) Suppose U is open with respect to $\|-\|_2$, i.e. for all $x \in U$, there exists an $r_x > 0$ such that $B_2(x, r_x) = \{y \in V \mid \|x - y\|_2 < r_x\} \subset U$.

Since $\|-\|_1$ and $\|-\|_2$ are strongly equivalent, there exists $D > 0$ such that $\|v\|_2 \leq D\|v\|_1$ for all $v \in V$. Let $x \in U$ and let $y \in B_1\left(x, \frac{1}{D}r_x\right) = \left\{z \in V \mid \|x - z\|_1 < \frac{1}{D}r_x\right\}$. Then

$$\frac{1}{D}\|x - y\|_2 \leq \|x - y\|_1 < \frac{1}{D}r_x.$$

We hence have that $\|x - y\|_2 < r_x$, so that $y \in B_2(x, r_x)$. Hence, $B_1\left(x, \frac{1}{D}r_x\right) \subset B_2(x, r_x) \subset U$, which proves that U is open with respect to $\|-\|_1$. $\qquad\square$

Naturally, since a normed space enjoys the same set of topological properties regardless of which norm is used (as long as they are all strongly equivalent), one can expect that sequences would behave similarly in any norm. In particular, if a sequence is either Cauchy or convergent in any one norm, then it will be Cauchy or convergent (with the same limit) in any other strongly equivalent norm.

Proposition 1.9.4. *Let V be a vector space and let $\|-\|_1$ and $\|-\|_2$ be strongly equivalent norms on V. Let $(x_n)_{n=1}^{\infty}$ be a sequence in V. Then:*

(a) *$(x_n)_{n=1}^{\infty}$ is Cauchy with respect to $\|-\|_1$ if and only if $(x_n)_{n=1}^{\infty}$ is Cauchy with respect to $\|-\|_2$.*

(b) *$\|x_n - x\|_1 \to 0$ if and only if $\|x_n - x\|_2 \to 0$.*

Proof.   (a) Firstly, since $\|-\|_1$ and $\|-\|_2$ are strongly equivalent norms, there exist $C, D > 0$ such that the following inequality holds for all $m, n \in \mathbb{N}$

$$0 \leq C\|x_m - x_n\|_1 \leq \|x_m - x_n\|_2 \leq D\|x_m - x_n\|_1.$$

($\implies$) Suppose $(x_n)_{n=1}^{\infty}$ is Cauchy with respect to $\|-\|_1$. Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that for all $m, n \geq N$, $\|x_m - x_n\|_1 < \epsilon/D$. Then for all $m, n \geq N$, we have

$$\|x_m - x_n\|_2 \leq D\|x_m - x_n\|_1 < D \times \frac{\epsilon}{D} = \epsilon.$$

($\impliedby$) Suppose $(x_n)_{n=1}^{\infty}$ is Cauchy with respect to $\|-\|_2$. Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that for all $m, n \geq N$, $\|x_m - x_n\|_2 < C\epsilon$. Then for all $m, n \geq N$, we have

$$\|x_m - x_n\|_1 \leq \frac{1}{C}\|x_m - x_n\|_2 < \frac{C\epsilon}{C} = \epsilon.$$

(b) For all $n \in \mathbb{N}$, we have that

$$0 \leq C\|x_n - x\|_1 \leq \|x_n - x\|_2 \leq D\|x_n - x\|_1$$

for some $C, D > 0$, since $\|-\|_1$ and $\|-\|_2$ are strongly equivalent norms.

Suppose $\lim_{n \to \infty} \|x_n - x\|_1 = 0$. Then

$$0 \leq \lim_{n \to \infty} \|x_n - x\|_2 \leq D \lim_{n \to \infty} \|x_n - x\|_1$$
$$0 \leq \lim_{n \to \infty} \|x_n - x\|_2 \leq 0.$$

This implies that $\|x_n - x\|_2 \to 0$ as $n \to \infty$ by the Squeeze Theorem. Similarly, if $\lim_{n \to \infty} \|x_n - x\|_2 = 0$, then

$$0 \leq \lim_{n \to \infty} \|x_n - x\|_1 \leq \frac{1}{C} \lim_{n \to \infty} \|x_n - x\|_2$$
$$0 \leq \lim_{n \to \infty} \|x_n - x\|_1 \leq 0$$

which implies that $\|x_n - x\|_1 \to 0$ as $n \to \infty$.

$\square$

This directly gives us the following result:

Proposition 1.9.5. *Let V be a vector space and let $\|-\|_1$ and $\|-\|_2$ be strongly equivalent norms on V. Then V is complete with respect to $\|-\|_1$ if and only if V is complete with respect to $\|-\|_2$.*

Proof. Suppose V is complete with respect to $\|-\|_1$. Let $(x_n)_{n=1}^{\infty}$ be a Cauchy sequence with respect to $\|-\|_2$. By Proposition 1.9.4 part (a), $(x_n)_{n=1}^{\infty}$ is Cauchy with respect to $\|-\|_1$. By completeness of V in the 1-norm, we have that $\|x_n - x\|_1 \to 0$ for some $x \in V$. But by Proposition 1.9.4 part (b), this implies that $\|x_n - x\|_2 \to 0$ as $n \to \infty$. This shows that every Cauchy sequence in $(V, \|-\|_2)$ has a limit, thus proving completeness. The same argument works in reverse for proving completeness of $(V, \|-\|_1)$ from that of $(V, \|-\|_2)$. $\square$

Finally, given a function $f : V \to W$ between finite dimensional normed spaces, the property of f being continuous is, in effect, independent on the particular norms used for V and W.

Proposition 1.9.6. *Let V, W be finite dimensional normed vector spaces. Then the continuity of a mapping $f : V \to W$ does not depend on the choice of norm on V and W.*

Proof. Suppose the map $f : V \to W$ is continuous with respect to the norm $\|-\|_1$ on V and $\|-\|_2$ on W. Let $\|-\|_{1'}$ be an arbitrary on V and $\|-\|_{2'}$ be an arbitrary norm on W. By Theorem 1.9.2, $\|-\|_1, \|-\|_{1'}$ are strongly equivalent norms on V and $\|-\|_2, \|-\|_{2'}$ are strongly equivalent norms on W. We will show that f is continuous with respect to the norms $\|-\|_{1'}$ and $\|-\|_{2'}$ on V and W, respectively.

Let $(x_n)_{n=1}^\infty$ be a sequence in V which converges to $x \in V$ in the norm $\|-\|_{1'}$. By Proposition 1.9.4, $x_n \to x$ as $n \to \infty$ with respect to $\|-\|_1$. By continuity of f (with respect to the unprimed norms) we have that $f(x_n) \to f(x)$ as $n \to \infty$ with respect to $\|-\|_2$. Another application of Proposition 1.9.4 shows that $f(x_n) \to f(x)$ as $n \to \infty$ with respect to $\|-\|_{2'}$. This shows the sequential continuity of f with respect to the primed norms, and hence demonstrates continuity. $\qquad\square$

In summary, we have shown that when we are working in Euclidean space E, we can equip E with any norm we want, and any topological result we deduce (including convergence and continuity) will be universal (i.e. norm independent). Thus, we do not necessarily need to work in the Euclidean norm $\|-\|_2$ all of the time, but rather we can choose convenient norms that are better suited for the particular problems we are working on at the time.

In particular, for each $p \in \mathbb{N}$, we have the following family of norms on E $\|-\|_p : E \to \mathbb{R}$

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

where the coefficients x_i are the coordinates of the vector $x \in E$ with respect to some basis of E. The norm $\|-\|_p$ is called the *p-norm* on E. If $p = 2$, the resulting p-norm is exactly the Euclidean norm $\|-\|_2$, which justifies the notation.

Another important norm we can define on E is the *maximum norm*, denoted $\|-\|_\infty$, which is defined by

$$\|x\|_\infty = \max_{i \in \{1,\ldots,n\}} \{|x_i|\}.$$

This is a very convenient norm to work with which we will frequently use, especially once we begin to do calculus on Euclidean space.

1.10 The Space $L(E,F)$

The importance of linear maps in multivariable calculus cannot be understated (indeed, we will later see that our definition of the derivative of a multivariable function will be an appropriate linear map). We will therefore carry out a study on the collection of all linear maps between two Euclidean spaces, and consider the structures that can be naturally prescribed to this space of maps to see what existing theories can be used to describe these objects and their relationships with one another.

Definition 1.10.1. Let E, F be vector spaces. Define $L(E, F)$ as the set of all linear maps $T : E \to F$.

Here, we will study the space $L(E, F)$ for the case where E and F are both Euclidean spaces (of possibly different dimension).

First, we formally introduce the *matrix representation* of an element of $L(E, F)$, which is a convenient way to describe and work with linear maps. In this representation, all of the information about a given linear map is captured by a (finite) array of real numbers. The details are as follows:

Let E be Euclidean n-space, F be Euclidean m-space and let $A \in L(E, F)$. Given a basis $\{e_1, \ldots, e_n\}$ for E, and a basis $\{\bar{e}_1, \ldots, \bar{e}_m\}$ for F, we can represent the linear map A as an $m \times n$ *matrix* with respect

to these bases, i.e. a collection of mn real numbers arranged in an $m \times n$ array (m rows and n columns). To see this, given any $x \in E$, we can express x in terms of the basis vectors

$$x = \sum_{j=1}^{n} x_j e_j.$$

We act on x with the linear map A, and by linearity we obtain

$$Ax = A\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j A e_j.$$

We can expand each $A e_j \in F$ in terms of the basis in F

$$A e_j = \sum_{i=1}^{m} A_{ij} \bar{e}_i.$$

Substituting this into our expression for Ax yields

$$Ax = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} x_j \bar{e}_i.$$

The constants A_{ij} are referred to as the entries of the matrix representing the linear map A. Notice that the entries depend on the choice of basis on both E and F.

The composition $B \circ A$ of two linear maps $A \in L(E, F)$ and $B \in L(F, G)$ is again a linear map. What will the corresponding matrix representation of $B \circ A$ look like in terms of the matrices for A and B? We'll use the same bases for E and F and consider the basis $\{\hat{e}_1, \ldots, \hat{e}_l\}$ for G (l-dimensional Euclidean space). Then for any $x \in E$, we can write

$$B \circ Ax = B\left(\sum_{j=1}^{m} \sum_{k=1}^{n} A_{jk} x_k \bar{e}_j\right)$$
$$= \sum_{j=1}^{m} \sum_{k=1}^{n} A_{jk} x_k B \bar{e}_j$$

where the final equality follows from the linearity of B. We expand $B\bar{e}_j$ in terms of the given basis in G, noting that the corresponding coefficients are simply the matrix elements B_{ij} of B

$$B \circ Ax = \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} B_{ij} A_{jk} x_k \hat{e}_i = \sum_{i=1}^{l} \sum_{k=1}^{n} C_{ik} x_k \hat{e}_i.$$

Hence, the linear map $B \circ A : E \to G$ is represented by an $l \times n$ matrix, and its corresponding matrix elements are C_{ik}, which are given by

$$C_{ik} = \sum_{j=1}^{m} B_{ij} A_{jk}.$$

One may recognise that the the matrix for the composite map $B \circ A$ is obtained by performing matrix multiplication with the matrices for B and A.

We now consider the relevant structures which we can endow upon $L(E, F)$. Firstly, from the algebraic properties of addition and scalar multiplication of vector-valued functions, we can consider $L(E, F)$ to be a vector space.

Theorem 1.10.2. *Let E, F be vector spaces with $\dim E = n$ and $\dim F = m$. Then $L(E, F)$ is an mn-dimensional vector space.*

Proof. Since sums of linear maps are linear, and scalar multiples of linear maps are linear, it follows that $L(E, F)$ is a vector space (a subspace of the vector space of all functions $f : E \to F$). We compute the dimension of $L(E, F)$ by explicitly finding a basis. The construction that will be employed below is a generalisation of the construction of the dual basis for E^* in Section 1.4. Fix a basis $\{e_1, \ldots, e_n\}$ for E and a basis $\{\bar{e}_1, \ldots, \bar{e}_m\}$ for F, and define linear maps $\phi_{ij} : E \to F$ for each $i = 1, \ldots, n$ and $j = 1, \ldots, m$ by

$$\phi_{ij}(e_k) = \delta_{jk}\bar{e}_i.$$

Note that defining the map ϕ_{ij} on the basis elements of E completely defines the map on all of E by linearity. We will now show that:

(i) The set $\{\phi_{ij} : E \to F \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ is linearly independent.

(ii) The span of $\{\phi_{ij} : E \to F \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ is equal to $L(E, F)$.

This will demonstrate that $\{\phi_{ij} : E \to F \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ is a basis for $L(E, F)$.

(i) Consider the following operator equation

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\phi_{ij} = 0$$

where the c_{ij} are some real constants, and the right-hand side is understood to be the zero operator in $L(E, F)$. Consider the action of the above operators on the basis vector e_k for an arbitrary $k \in \{1, \ldots, n\}$

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\phi_{ij}(e_k) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\delta_{jk}\bar{e}_i = \sum_{j=1}^{n} c_{ik}\bar{e}_i = 0.$$

However, by linear independence of $\{\bar{e}_1, \ldots, \bar{e}_m\}$, it follows that $c_{ik} = 0$ for all $i \in \{1, \ldots, m\}$. This holds for any $k \in \{1, \ldots, n\}$, which implies that the only solution to the initial operator equation is the trivial solution. This proves linear independence.

(ii) Let $A \in L(E, F)$, and let $x \in E$. From a previous calculation, we have that

$$\begin{aligned}
Ax &= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}x_j\bar{e}_i \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}\sum_{k=1}^{n} x_k\phi_{ij}(e_k) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}\phi_{ij}\left(\sum_{k=1}^{n} x_k e_k\right) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}\phi_{ij}(x).
\end{aligned}$$

Since the above holds for all $x \in E$, we must have equality of the operators

$$A = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}\phi_{ij}.$$

So A can be written as a linear combination of the ϕ_{ij}. Moreover, the coefficients are none other than the entries A_{ij} of the matrix corresponding to A. In fact, one sees that the matrix representation of the linear map ϕ_{ij} simply contains a '1' in the ith row and j^{th} column, and zeros elsewhere.

Now, since the set $\{\phi_{ij} : E \to F \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ is a basis for $L(E, F)$, the dimension of $L(E, F)$ is simply the number of basis elements. There are $m \times n = mn$ such linear maps in this basis, and so it follows that $\dim L(E, F) = mn$. $\qquad\square$

The dual space E^* of Euclidean n-space E can be alternatively expressed as $L(E, \mathbb{R})$. We know that the dimension of E^* is n - the same as E. This is consistent with the above theorem, which says that $\dim E^* = \dim L(E, \mathbb{R}) = n \times 1 = n$.

Next, we will see that $L(E, F)$ inherits a normed structure from E and F, and so can be considered to be a normed vector space. Thus, all of our work on studying finite dimensional normed vector spaces can be applied to $L(E, F)$.

Theorem 1.10.3. *Let E, F be finite dimensional normed vector spaces. Then $L(E, F)$ is a normed vector space, with norm given by*

$$\|T\| = \sup_{\|x\|_E = 1} \{\|Tx\|_F \mid x \in E\}$$

where $\|-\|_E$ is the norm on E and $\|-\|_F$ is the norm on F.

Proof. Let $T \in L(E, F)$. Denote by S the unit sphere in E with respect to the norm on E, i.e.

$$S = \{x \in E \mid \|x\|_E = 1\}.$$

The set S is a compact subset of E. Since T is linear, it is continuous by Theorem 1.7.6. Thus, the composition of the norm $\|-\|_F$ with T is a continuous function. Hence, by the Extreme Value Theorem, the restriction of $\|-\|_F \circ T$ to S attains a maximum value, which is defined to be $\|T\|$. This shows that the function $\|-\| : L(E, F) \to \mathbb{R}$ described above has well-defined values on all of $L(E, F)$. We will proceed to show that this function is indeed a norm on $L(E, F)$.

(i) (Positivity) Let $T \in L(E, F)$. Then $\|Tx\|_F \geq 0$ for all $x \in S$ by positivity of the norm $\|-\|_F$. Taking the supremum over S on both sides gives $\|T\| = \sup_{x \in S} \|Tx\|_F \geq 0$.

If $T = 0$ (i.e. the zero function), then $\|Tx\|_F = 0$ for all $x \in S$, and so $\|T\| = \sup_{x \in S} \|Tx\|_F = 0$. Conversely, if $\|T\| = 0$, then for all $x \in S$, we have

$$0 \leq \|Tx\|_F \leq \sup_{x \in S} \|Tx\|_F = \|T\| = 0.$$

So $\|Tx\|_F = 0$ for all $x \in S$. Let $y \in E \backslash \{0\}$. Then $y/\|y\|_E \in S$, so that

$$0 = \left\| T \frac{y}{\|y\|_E} \right\|_F = \frac{1}{\|y\|_E} \|Ty\|_F.$$

Thus, $\|Ty\|_F = 0$, which implies that $Ty = 0$. Thus, T maps every element of E to $0 \in F$, and so we must have that $T = 0$.

(ii) (Homogeneity of degree 1) Let $a \in \mathbb{R}$. Then

$$\|aT\| = \sup_{x \in S} \|aTx\|_F = |a| \sup_{x \in S} \|Tx\|_F = |a| \|T\|.$$

(iii) (Triangle Inequality) Let $A, B \in L(E, F)$. Then

$$\|A + B\| = \sup_{x \in S} \|(A + B)x\|_F \leq \sup_{x \in S} (\|Ax\|_F + \|Bx\|_F) = \|A\| + \|B\|.$$

$\qquad\square$

The norm defined on $L(E,F)$ is often referred to as the 'operator norm'. It is a natural norm induced by the norms on E and F. A geometric interpretation for the norm of a linear map T is the maximum extent/scale factor that a unit vector is scaled by T. So the value of $\|Tx\|$ never 'blows up' as we approach any $x_0 \in E$ with finite length. This is a feature of the finite dimensional nature of the spaces we are working in.

Linear operators can be seen to satisfy the following boundedness property. Given any $x \in E \backslash \{0\}$, we have

$$\|Tx\| = \|x\|\|T\hat{x}\|$$
$$\leq \|x\| \sup_{\|y\|=1} \|Ty\|$$
$$= \|T\|\|x\|$$

where $\hat{x} = x/\|x\|$ is the vector x normalised to have unit length. This inequality is also valid for $x = 0$, in which case we have the trivial equality $0 = 0$, and so the inequality holds for all $x \in E$. In general, an operator satisfying this inequality is called a *bounded operator*. The name arises from the fact that a bounded operator maps bounded sets to bounded sets. In the finite dimensional case, every linear operator is bounded.

The value of the operator norm will depend on which norms we use for E and F, but since $L(E,F)$ is a finite dimensional vector space, every possible choice will yield strongly equivalent norms, and so it doesn't really matter too much which norms on E and F we pick for the purposes of doing analysis.

Finally, we will present some further results on linear maps. In particular, we will present a version of the First Isomorphism Theorem for vector spaces, and from it derive the famous 'Rank Theorem'. We then show a situation in which the injectivity of a linear map becomes a sufficient property for bijectivity.

In the results that follow, we will make use of the notion of a *quotient space*. A somewhat geometrical picture is the following: given a vector space V and a subspace W of V, we wish to consider the space W within V (which contains the origin), and all possible translations of W by vectors in V. The set consisting of W as well as all of its different translates is what we will call the quotient space V/W. So elements of V/W are subsets of V, called cosets. When are two vectors $v_1, v_2 \in V$ contained in the same coset? Well, if they were, then we could write each vector as the sum of some vector in W and the constant offset/translation vector $t \in V$, i.e.

$$v_1 = w_1 + t$$
$$v_2 = w_2 + t$$

for some $w_1, w_2 \in W$. Hence, the difference between v_1 and v_2 is

$$v_1 - v_2 = (w_1 + t) - (w_2 + t) = w_1 - w_2 \in W$$

where the statement $w_1 - w_2 \in W$ follows from the fact that W is a subspace. This motivates us to define the relation $\sim$ on V by

$$v_1 \sim v_2 \quad \text{if } v_1 - v_2 \in W.$$

This is an equivalence relation on V, which follows from the fact that W is a subspace. The equivalence classes $[v]$ of V under the relation $\sim$ are the vector cosets $v + W$ (translation of W by v). To show this, let $v' \in [v]$. Then $v' - v := w \in W$, and so $v' = v + (v' - v) = v + w$, which shows that $v' \in v + W$, implying that $[v] \subseteq v + W$. Conversely, if we suppose that $v' \in v + W$, then $v' = v + w$ for some $w \in W$. This implies that $v' - v = w \in W$, which shows that $v' \sim v$, i.e. $v' \in [v]$ and hence $v + W \subseteq [v]$. Together, we have that $[v] = v + W$. This leads us to the following definition:

Definition 1.10.4. Let V be a vector space, and let W be a subspace of V. The **quotient space** V/W is the set of vector cosets

$$V/W = \{v + W \mid v \in V\}.$$

We can define addition and scalar multiplication on V/W by $(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W$ and $a(v_1 + W) = av_1 + W$. These definitions are appropriate, since we have:

Proposition 1.10.5. *Let V be a vector space, and let W be a subspace of V. The operations of addition and scalar multiplication of vector cosets are well-defined operations on the quotient space V/W.*

Proof. Let $v_1, v_2, v_1', v_2' \in V$ such that

$$v_1 + W = v_1' + W$$
$$v_2 + W = v_2' + W.$$

This implies that $v_1 \sim v_1'$ and $v_2 \sim v_2'$, and hence that $v_1' - v_1 \in W$ and $v_2' - v_2 \in W$. We therefore have that

$$(v_1' + W) + (v_2' + W) = (v_1' - v_1 + v_1 + W) + (v_2' - v_2 + v_2 + W)$$
$$= (v_1 + W) + (v_2 + W).$$

Now let $a \in \mathbb{R}$. Then

$$av_1' + W = a(v_1' + W) = a(v_1' - v_1 + v_1 + W) = a(v_1 + W) = av_1 + W.$$

$\square$

Thus, when equipped with these operations, V/W is a vector space. The fact that V/W satisfies the defining properties of a vector space follows quite readily from simple and not very illuminating calculations involving the analogous properties of V. The additive identity on V/W is $0 + W = W$.

We now have the machinery required to present the (coincidentally) first major theorem for this section: the First Isomorphism Theorem (for vector spaces), which states that given a linear map $T : V \to W$ between vector spaces, the quotient space $V/\ker T$ is isomorphic to the range $\mathcal{R}(T)$ of T. This, at first glance, may seem like magic, but actually once you break it down it really is just common sense! Let's explain:

A desired isomorphism $\widetilde{T} : V/\ker T \to \mathcal{R}(T)$ will be constructed from T. How does T itself fail to be an isomorphism? Since it is a vector space homomorphism (i.e. a linear map), the only way it can fall short of being an isomorphism is if it fails to be bijective. We form the map $\widetilde{T}$ by 'patching in' the deficiencies in T by 'fixing' injectivity and surjectivity. It turns out that the process of fixing these failures shrinks the codomain to the image of T and shrinks the domain to the quotient space $V/\ker T$.

Theorem 1.10.6. *(First Isomorphism Theorem) Let V, W be vector spaces and let $T : V \to W$ be a linear map. Then $V/\ker T \simeq \mathcal{R}(T)$.*

Proof. The map T can be made surjective by restricting its codomain to its range $\mathcal{R}(T)$, so that $T : V \to \mathcal{R}(T)$ is a surjective linear map.

A linear map T is injective if and only if $\ker T = \{0\}$. So if T is not injective, then its kernel is a subspace of V whose dimension is nonzero. In order to 'shrink' the kernel, we form the quotient space $V/\ker T$, which in essence collapses all vectors in the kernel down to a single element; a vector coset. We hence define a map $\widetilde{T} : V/\ker T \to \mathcal{R}(T)$ that acts on the cosets by

$$\widetilde{T}(v + \ker T) = T(v).$$

We will need to show that

(i) $\widetilde{T}$ is well-defined.

(ii) $\widetilde{T}$ is linear.

(iii) $\widetilde{T}$ is injective.

(iv) $\widetilde{T}$ is surjective.

Together, these would show that $\widetilde{T}$ is an isomorphism between $V/\ker T$ and $\mathcal{R}(T)$.

(i) Let $v' \in V$ be such that $v' + \ker T = v + \ker T$, i.e. $v' \sim v$. Thus, $v' - v \in \ker T$. We therefore have that

$$\widetilde{T}(v' + \ker T) = T(v') = T(v' - v + v) = T(v' - v) + T(v) = 0 + T(v) = T(v) = \widetilde{T}(v + \ker T).$$

(ii) Let $v_1 + \ker T, v_2 + \ker T \in V/\ker T$. Then

$$\begin{aligned}
\widetilde{T}(v_1 + \ker T + v_2 + \ker T) &= \widetilde{T}(v_1 + v_2 + \ker T) \\
&= T(v_1 + v_2) \\
&= T(v_1) + T(v_2) \\
&= \widetilde{T}(v_1 + \ker T) + \widetilde{T}(v_2 + \ker T).
\end{aligned}$$

Let $a \in \mathbb{R}$ and let $v + \ker T \in V/\ker T$. Then

$$\begin{aligned}
\widetilde{T}(a(v + \ker T)) &= \widetilde{T}(av + \ker T) \\
&= T(av) \\
&= aT(v) \\
&= a\widetilde{T}(v + \ker T).
\end{aligned}$$

(iii) Let $(v + \ker T) \in \ker \widetilde{T}$. So $\widetilde{T}(v + \ker T) = 0$. But

$$0 = \widetilde{T}(v + \ker T) = T(v).$$

This implies that $v \in \ker T$, and hence $v + \ker T = \ker T$. Thus, $\ker \widetilde{T} = \{\ker T\}$, i.e. the kernel of $\widetilde{T}$ contains only the zero coset. Since $\widetilde{T}$ is linear, this proves the injectivity of $\widetilde{T}$.

(iv) Let $w \in \mathcal{R}(T)$. Since $T : V \to \mathcal{R}(T)$ is surjective, there exists a $v \in V$ such that $T(v) = w$. But then we have that $v + \ker T \in V/\ker T$ satisfies

$$\widetilde{T}(v + \ker T) = T(v) = w.$$

<div align="right">□</div>

This line of thinking can be used to prove other versions of the First Isomorphism Theorem between different structures, e.g. groups, rings, fields. The central idea remains the same: given a map that is *not* bijective, how can we 'fix' it and make it bijective? We fix surjectivity by restricting the codomain to the image of the map, and fix injectivity by grouping together elements that map to the same object. Of course, this means that the modified map now acts on *equivalence classes* rather than the original elements. Finally, if the map preserves some structure (e.g. is a group homomorphism), then the modified map will also preserve the same structure (e.g. group multiplication).

Back to quotient spaces: what is the dimension of the quotient space? Given that in the construction of the quotient space, we group together entire subsets of vectors to form a single vector coset, we expect that the quotient space is 'smaller' than the original vector space. Exactly how much smaller is dictated by the size of the subspace that has been quotiented out.

Proposition 1.10.7. *Let V be a finite dimensional vector space, and let W be a subspace of V. Then* $\dim V/W = \dim V - \dim W$.

Proof. Suppose $\dim V = n$, $\dim W = k \leq n$, $\{w_1, \ldots, w_k\}$ is a basis for W and $\{v_1, \ldots, v_n\}$ be a basis for V. By part (a) of the Replacement Lemma (Lemma 1.1.4) and Corollary 1.1.8, we can replace k of the basis vectors for V with the basis vectors for W and obtain (after some relabelling) the following basis for V:

$$\{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\}.$$

We claim that $\{v_{k+1} + W, \ldots, v_n + W\}$ is a basis for V/W.

Let $v + W \in V/W$. We can expand $v \in V$ as a linear combination of our basis for V

$$v = c_1 w_1 + \cdots + c_k w_k + c_{k+1} v_{k+1} + \cdots + c_n v_n.$$

This implies that

$$v + W = c_1(w_1 + W) + \cdots + c_k(w_k + W) + c_{k+1}(v_{k+1} + W) + \cdots + c_n(v_n + W).$$

But $w_i + W = W$ for each $i \in \{1, \ldots, k\}$, so that

$$v + W = c_{k+1}(v_{k+1} + W) + \cdots + c_n(v_n + W).$$

Thus, any element of V/W can be expressed as a linear combination of the vectors in

$$\{v_{k+1} + W, \ldots, v_n + W\}.$$

This shows that

$$V/W = \operatorname{span}\{v_{k+1} + W, \ldots, v_n + W\}.$$

Next, consider the following equation

$$a_1(v_{k+1} + W) + \cdots + a_{n-k}(v_n + W) = W$$

where $a_1, \ldots, a_{n-k} \in \mathbb{R}$. Performing the addition of the vector cosets allows us to write

$$(a_1 v_{k+1} + \cdots + a_{n-k} v_n) + W = W,$$

which is satisfied if and only if

$$a_1 v_{k+1} + \cdots + a_{n-k} v_n \in W.$$

Thus, the above vector can be written as a linear combination of the basis vectors for W

$$a_1 v_{k+1} + \cdots + a_{n-k} v_n = b_1 w_1 + \cdots + b_k w_k.$$

We can rearrange this as follows

$$a_1 v_{k+1} + \cdots + a_{n-k} v_n - b_1 w_1 - \cdots - b_k w_k = 0.$$

But since $\{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\}$ is a basis for V, it is a linearly independent set. All of the above coefficients must vanish, although in particular, we have that $a_1 = \cdots = a_{n-k} = 0$, which implies the linear independence of $\{v_{k+1} + W, \ldots, v_n + W\}$, and thus demonstrates that it is a basis for V/W. The number of elements in this basis is equal to $n - k$, so we have that

$$\dim(V/W) = n - k = \dim V - \dim W.$$

$\square$

The above result, together with the First Isomorphism Theorem, leads to an easy proof of the Rank Theorem, which provides a relationship between the kernel and range of a linear map.

Theorem 1.10.8. *(Rank Theorem) Let V, W be finite dimensional vector spaces and let $T : V \to W$ be a linear map. Then*

$$\dim \ker T + \dim \mathcal{R}(T) = \dim V.$$

Proof. By the First Isomorphism Theorem, $V/\ker T$ is isomorphic to $\mathcal{R}(T)$. By Theorem 1.1.11, this implies that $\dim(V/\ker T) = \dim \mathcal{R}(T)$. But by Proposition 1.10.7, we have that

$$\dim \mathcal{R}(T) = \dim(V/\ker T) = \dim V - \dim \ker T.$$

Rearranging the above expression gives the desired relationship. $\square$

The name 'Rank Theorem' is due to the fact that the dimension of the range of T is often referred to as the *rank* of the linear operator T.

Finally, we provide a quick application of the Rank Theorem:

Proposition 1.10.9. *Let E and F be n-dimensional vector spaces, and let $T \in L(E, F)$. Then T is injective if and only if T is bijective (and hence an isomorphism).*

Proof. Suppose T is injective. Since T is linear, this implies that $\ker T = \{0\}$, so that $\dim \ker T = 0$. Then by the Rank Theorem, we have that

$$\dim \mathcal{R}(T) = \dim E = n = \dim F.$$

However, since $\mathcal{R}(T) \subseteq F$, we must have that $\mathcal{R}(T) = F$, as the only n-dimensional subspace of an n-dimensional vector space F is F itself. This shows that T is surjective, and hence a bijection. $\square$

This result will be useful to establish when a linear map is *invertible* (recall that a function is invertible if and only if it is bijective). For instance, this will pop up when we come to discuss the *Inverse Function Theorem*, which is concerned with the inverse of a differentiable function.

There are many more purely algebraic and topological things in Euclidean space which we can discuss further, but for now, we'll call it a day and start doing some calculus!

Chapter 2

Differentiation

2.1 Definition of the Derivative

We will now begin developing the theory of calculus for vector-valued functions with vector-valued inputs, starting with *differential calculus*. When talking about the differentiability of mappings between Euclidean spaces, we will consider the domain of the map to be an open set, so that we can move some small amount in any direction about any point in the domain. This, loosely speaking, will give us some notion of being able to approach any particular point from any given direction, and thus allow us to consider the rate of change of the function along this path. Indeed, the rate of change of a function as we vary the input variable(s) is one intuitive notion of the derivative of the function.

In the discussion that will follow, let E be Euclidean n-space, let F be Euclidean m-space and let $U \subseteq E$ be an open set. We want to define the notion of differentiability for mappings $f : U \to F$. Another way to think whether a map f is differentiable is whether it looks locally like a first-order polynomial, i.e. in some neighbourhood of a point x_0 in the domain of f, the following expression is a 'good' approximation for f

$$f(x) \approx f(x_0) + T(x - x_0)$$

for some linear map $T \in L(E, F)$. Of course, the behaviour of f can be quite complicated, and so such a linear approximation for f may be quite bad far away from x_0. The expression is of course exact *at* x_0. As for points 'close to' x_0, we will designate the functions f for which the approximation is 'good' to be the differentiable functions. The linear map T appearing in the linear approximation (the 'slope', if you may, to make an analogy with real-valued functions of a single variable) that minimises the error of the linear approximation is what we will call the derivative of f, so that the derivative is the linear map used to construct the 'best' linear approximation of f at x_0.

Definition 2.1.1. A mapping $f : U \to F$ is said to be **differentiable** at $x_0 \in U$ if there exists a linear map $Df(x_0) \in L(E, F)$ and a map $R : U \to F$ such that for every $x \in U$, we have

$$f(x) = f(x_0) + [Df(x_0)](x - x_0) + R(x)$$

and

$$\lim_{\|x - x_0\| \to 0} \frac{\|R(x)\|}{\|x - x_0\|} = 0.$$

If f is differentiable at x_0, the linear map $Df(x_0) : E \to F$ is called the **derivative** of f at x_0.

So, we can express any function f as the sum of a linear approximation for f and some remainder/rubbish term $R(x)$, the latter of which captures the 'error' of the linear approximation compared to the true value of f. The value of the error can be quite large for certain values of x, but if f is differentiable, the error tends to zero as one approaches x_0. This makes rigourous the notion that the linear approximation can be made arbitrarily accurate for values of x that are sufficiently close to x_0.

Since the derivative is a linear map (i.e. an element of $L(E, F)$), it can be represented by an $m \times n$ matrix. This matrix is called the *Jacobian matrix*.

If f is differentiable at each $x_0 \in U$, we say that f is differentiable on U (or simply: differentiable), and can define the differentiation map $\mathrm{D}f : U \to L(E, F)$, which takes each $x_0 \in U$ to $\mathrm{D}f(x_0)$; the derivative of f at x_0.

Let's look at a couple of examples. First, consider the first-order polynomial $p : \mathbb{R}^n \to \mathbb{R}^m$

$$p(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

where $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ is a linear map and $\mathbf{b} \in \mathbb{R}^m$ is a constant vector. We claim that p is a differentiable function for all $\mathbf{x_0} \in \mathbb{R}^n$ with derivative A. To see this, define the function $R : \mathbb{R}^n \to \mathbb{R}^m$ by

$$R(\mathbf{x}) = p(\mathbf{x}) - p(\mathbf{x_0}) - A(\mathbf{x} - \mathbf{x_0}).$$

Then substituting the definition of p gives

$$R(\mathbf{x}) = A\mathbf{x} + \mathbf{b} - (A\mathbf{x_0} + \mathbf{b}) - A\mathbf{x} + A\mathbf{x_0} = \mathbf{0}$$

which holds for all $\mathbf{x} \in \mathbb{R}^n$. Thus, we clearly have that

$$\lim_{\|\mathbf{x}-\mathbf{x_0}\| \to 0} \frac{\|R(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x_0}\|} = 0.$$

This shows that p is differentiable on $\mathbb{R}^n$, and moreover that $\mathrm{D}p(\mathbf{x_0}) = A$ for all $\mathbf{x_0} \in \mathbb{R}^n$, i.e. the differentiation map $\mathrm{D}p : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$ is constant with value A. Notice that the remainder function R is identically zero on $\mathbb{R}^n$. This is because p is a first-order polynomial in $\mathbf{x}$, so it is *equal* to its linear approximation. The Jacobian matrix for p at any $\mathbf{x_0} \in \mathbb{R}^n$ would simply be the matrix representation of A.

Now for a slightly more nontrivial example. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$f(x, y) = \begin{pmatrix} x^2 + y^2 \\ xy \end{pmatrix}$$

where we have expressed all elements of $\mathbb{R}^2$ in terms of the standard basis. We will show that f is differentiable on $\mathbb{R}^2$, with the differentiation map $\mathrm{D}f : \mathbb{R}^2 \to L(\mathbb{R}^2, \mathbb{R}^2)$ given by

$$\mathrm{D}f(x, y) = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}.$$

Let $\mathbf{x_0} = (x_0, y_0)^T \in \mathbb{R}^2$. Using the above linear map as our candidate for the derivative, we will define the remainder function $R : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$R(x, y) = f(x, y) - f(x_0, y_0) - [\mathrm{D}f(x_0)](\mathbf{x} - \mathbf{x_0}).$$

Writing the above expression out as a vector equation gives

$$\begin{aligned} R(x, y) &= \begin{pmatrix} x^2 + y^2 \\ xy \end{pmatrix} - \begin{pmatrix} x_0^2 + y_0^2 \\ x_0 y_0 \end{pmatrix} - \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \\ &= \begin{pmatrix} x^2 - x_0^2 + y^2 - y_0^2 \\ xy - x_0 y_0 \end{pmatrix} - \begin{pmatrix} 2x(x - x_0) + 2y(y - y_0) \\ y(x - x_0) + x(y - y_0) \end{pmatrix} \\ &= \begin{pmatrix} -(x - x_0)^2 - (y - y_0)^2 \\ -(x - x_0)(y - y_0) \end{pmatrix}. \end{aligned}$$

Consider taking the maximum norm of both sides

$$\|R(x, y)\|_\infty = \max \left\{ \left| (x - x_0)^2 + (y - y_0)^2 \right|, \left| (x - x_0)(y - y_0) \right| \right\}.$$

But since $\|\mathbf{x} - \mathbf{x_0}\|_\infty = \max\{|x - x_0|, |y - y_0|\}$, we have that

$$\|R(x, y)\|_\infty \leq \max\left\{2\|\mathbf{x} - \mathbf{x_0}\|_\infty^2, \|\mathbf{x} - \mathbf{x_0}\|_\infty^2\right\} = 2\|\mathbf{x} - \mathbf{x_0}\|_\infty^2.$$

Dividing through by $\|\mathbf{x} - \mathbf{x_0}\|_\infty$ yields

$$0 \leq \frac{\|R(x, y)\|_\infty}{\|\mathbf{x} - \mathbf{x_0}\|_\infty} \leq 2\|\mathbf{x} - \mathbf{x_0}\|_\infty.$$

Taking the limit as $\|\mathbf{x} - \mathbf{x_0}\|_\infty \to 0$ and applying the Squeeze Theorem yields

$$\lim_{\|\mathbf{x} - \mathbf{x_0}\|_\infty \to 0} \frac{\|R(x, y)\|_\infty}{\|\mathbf{x} - \mathbf{x_0}\|_\infty} = 0.$$

Since the maximum norm is strongly equivalent to the Euclidean norm, the above limit also converges to 0 in the Euclidean norm. This shows that f is differentiable at $\mathbf{x_0}$. Since the choice of $\mathbf{x_0} \in \mathbb{R}^2$ was arbitrary, it follows that f is differentiable on $\mathbb{R}^2$, with derivative $\mathrm{D}f$ as shown.

Our use of the maximum norm over the Euclidean norm to prove the differentiability of f resulted in a greatly simplified calculation. Our argument for why this was justified may be generalised to the following result.

Proposition 2.1.2. *The differentiability of a mapping $f : U \to F$ does not depend on the choice of norm on E and F.*

Proof. Suppose $f : U \to F$ is differentiable at $x_0 \in U$ with respect to the norms $\|-\|_E$ and $\|-\|_F$ on E and F respectively. That is, there exist maps $\mathrm{D}f(x_0) \in L(E, F)$ and $R : U \to F$ such that for all $x \in U$,

$$f(x) = f(x_0) + [\mathrm{D}f(x_0)](x - x_0) + R(x)$$

and

$$\lim_{\|x - x_0\|_E \to 0} \frac{\|R(x)\|_F}{\|x - x_0\|_E} = 0.$$

Let $\|-\|_{E'}$ be any other norm on E and $\|-\|_{F'}$ be an arbitrary norm on F. Then by Theorem 1.9.2, $\|-\|_E$ and $\|-\|_{E'}$ are strongly equivalent, and $\|-\|_F$ and $\|-\|_{F'}$ are strongly equivalent. This means that there exists $K_1, K_2 > 0$ such that

$$\|x - x_0\|_E \leq K_1 \|x - x_0\|_{E'}$$
$$\|R(x)\|_{F'} \leq K_2 \|R(x)\|_F$$

which holds for all $x \in U$. Thus, we have that

$$0 \leq \frac{\|R(x)\|_{F'}}{\|x - x_0\|_{E'}} \leq K_1 K_2 \frac{\|R(x)\|_F}{\|x - x_0\|_E}.$$

Taking limits gives

$$0 \leq \lim_{\|x - x_0\|_{E'} \to 0} \frac{\|R(x)\|_{F'}}{\|x - x_0\|_{E'}} \leq K_1 K_2 \lim_{\|x - x_0\|_{E'} \to 0} \frac{\|R(x)\|_F}{\|x - x_0\|_E}.$$

But by Proposition 1.9.4, we have that $\|x - x_0\|_E \to 0$ if and only if $\|x - x_0\|_{E'} \to 0$. Thus, the limit on the right-hand side of the above inequality vanishes, and thus we conclude by the Squeeze Theorem that

$$\lim_{\|x - x_0\|_{E'} \to 0} \frac{\|R(x)\|_{F'}}{\|x - x_0\|_{E'}} = 0$$

which proves that f is differentiable at x_0 with respect to the norms $\|-\|_{E'}$ and $\|-\|_{F'}$, with the same derivative $\mathrm{D}f(x_0)$ and remainder function R. $\square$

You may have noticed a few things about the calculations in the previous examples. Firstly, they were somewhat tedious to do! Indeed, in the future we surely do not want to be showing that a given function is differentiable directly from the definition. Secondly (and perhaps most importantly), to show that a function is differentiable, we somehow have to pre-ordain what the derivative ought to be (which we did before by coincidentally materialising the correct linear map seemingly out of nowhere). These factors combined means that the definition of the derivative alone does not naturally lend itself to *practical* computations for the derivative of a multivariable function, and so we'll need to develop the theory further to obtain some results and formulae to simplify this process.

The definition of the derivative can actually be stated in terms of a single limit expression by combining the two expressions in Definition 2.1.1 by eliminating the remainder function R. Explicitly, a function $f : U \to F$ is differentiable at $x_0 \in U$ if and only if there exists a linear map $\mathrm{D}f(x_0) \in L(E, F)$ such that

$$\lim_{\|x - x_0\| \to 0} \frac{\|f(x) - f(x_0) - [\mathrm{D}f(x_0)](x - x_0)\|}{\|x - x_0\|} = 0.$$

The derivative of a function f at x_0 retains its interpretation as the infinitesimal rate of change of the function f as you move away from x_0 even in the multivariable realm. Given any vector $y \in E$, the derivative at x_0 evaluated at y, $[\mathrm{D}f(x_0)](y)$, is the variation of f as one moves an infinitesimal distance in a straight line in the direction of y from the point x_0.

Proposition 2.1.3. *Suppose $f : U \to F$ is differentiable at $x_0 \in U$. Then for every $y \in E$, we have that*

$$[Df(x_0)](y) = \lim_{t \to 0} \frac{f(x_0 + ty) - f(x_0)}{t}.$$

Proof. Let $y \in E$. If $y = 0$, then

$$\lim_{t \to 0} \frac{f(x_0 + ty) - f(x_0)}{t} = \lim_{t \to 0} \frac{f(x_0) - f(x_0)}{t} = 0.$$

Moreover, we have that $[\mathrm{D}f(x_0)](0) = 0$ since $\mathrm{D}f(x_0)$ is linear, which verifies the statement for this particular case.

Now suppose that $y \neq 0$. Since U is open, there exists a $\delta > 0$ such that $B(x_0, \delta) \subseteq U$. Then for all $t \in (-\delta/\|y\|, \delta/\|y\|)$, we have that $x = x_0 + ty \in U$, and hence

$$f(x_0 + ty) - f(x_0) = [\mathrm{D}f(x_0)](ty) + R(x)$$

since f is differentiable at x_0. Dividing through the above expression by t and rearranging gives

$$\frac{f(x_0 + ty) - f(x_0)}{t} - \frac{t[\mathrm{D}f(x_0)](y)}{t} = \frac{R(x)}{t}$$

noting that we have used the linearity of $\mathrm{D}f(x_0)$ to pull out the factor of t. Now, by applying the norm to both sides and taking the limit $t \to 0$ gives

$$\lim_{t \to 0} \left\| \frac{f(x_0 + ty) - f(x_0)}{t} - [\mathrm{D}f(x_0)](y) \right\| = \lim_{t \to 0} \frac{\|R(x)\|}{|t|}.$$

Since $\|x - x_0\| = \|ty\| = |t|\|y\|$, we can write

$$\left\| \lim_{t \to 0} \frac{f(x_0 + ty) - f(x_0)}{t} - [\mathrm{D}f(x_0)](y) \right\| = \|y\| \lim_{t \to 0} \frac{\|R(x)\|}{|t|\|y\|}$$

$$= \|y\| \lim_{\|x - x_0\| \to 0} \frac{\|R(x)\|}{\|x - x_0\|}$$

$$= 0$$

which follows from the differentiability of f at x_0 and by exploiting the continuity of the norm in two places: to assert that $\|x - x_0\| \to 0$ as $t \to 0$ and to pull limits inside of the norm. By positivity of the norm, we obtain the desired result

$$\lim_{t \to 0} \frac{f(x_0 + ty) - f(x_0)}{t} = [\mathrm{D}f(x_0)](y).$$

$\square$

This leads us to define the concept of the *directional derivative*, a weaker notion of the derivative of a vector-valued function.

Definition 2.1.4. The **directional derivative** of a map $f : U \to F$ at $x_0 \in U$ in the direction of $y \in E \backslash \{0\}$ is

$$\lim_{t \to 0} \frac{f(x_0 + t\hat{\mathbf{y}}) - f(x_0)}{t},$$

provided that the above limit exists. The vector $\hat{\mathbf{y}} = y/\|y\|$ is a unit vector in the direction of y.

An immediate consequence of Proposition 2.1.3 is that it provides a convenient relation between the derivative of a differentiable function and its corresponding directional derivatives.

Corollary 2.1.5. *Suppose $f : U \to F$ is differentiable at $x_0 \in U$. Then for every $y \in E \backslash \{0\}$, the directional derivative of f at x_0 in the direction of y exists, and is equal to $[Df(x_0)](\hat{\mathbf{y}})$.*

The converse of the above statement does not hold. That is, if the directional derivative of f at x_0 in the direction of y exists for every $y \in E$, f is not necessarily differentiable at x_0. A nice counterexample is the following: define the function $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x, y) = \begin{cases} \frac{x^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Let $\mathbf{v} \in \mathbb{R}^2$ be any nonzero vector. We can express $\hat{\mathbf{v}}$ in terms of the standard basis $\{\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y\}$ in $\mathbb{R}^2$ as follows

$$\hat{\mathbf{v}} = \cos \theta \hat{\mathbf{e}}_x + \sin \theta \hat{\mathbf{e}}_y$$

for some $\theta \in \mathbb{R}$. Then for any $t \neq 0$, we have

$$\lim_{t \to 0} \frac{f(\mathbf{0} + t\hat{\mathbf{v}}) - f(\mathbf{0})}{t} = \lim_{t \to 0} \frac{1}{t} \left(\frac{t^3 \cos^3 \theta}{t^2 \cos^2 \theta + t^2 \sin^2 \theta} \right) = \lim_{t \to 0} \frac{t^3 \cos^3 \theta}{t^3} = \lim_{t \to 0} \cos^3 \theta = \cos^3 \theta$$

noting that we have used the fact that $\cos^2 \theta + \sin^2 \theta = 1$. This shows that the directional derivative of f at $\mathbf{0}$ in the direction of $\mathbf{v}$ exists, and is equal to $\cos^3 \theta$. Since the choice of $\mathbf{v} \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$ was arbitrary, it follows that all of the directional derivatives of f at $\mathbf{0}$ exist. However, f is *not* differentiable at $\mathbf{0}$. Assume for a contradiction that f were differentiable at $\mathbf{0}$ with derivative $\mathrm{D}f(\mathbf{0})$. By Corollary 2.1.5, $[\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_x)$ is equal to the directional derivative of f at $\mathbf{0}$ in the direction of $\hat{\mathbf{e}}_x$, so that

$$[\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_x) = \cos^3 \theta \Big|_{\theta=0} = 1.$$

Similarly, $[\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_y)$ is equal to the directional derivative of f at $\mathbf{0}$ in the direction of $\hat{\mathbf{e}}_y$, which implies that

$$[\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_y) = \cos^3 \theta \Big|_{\theta=\frac{\pi}{2}} = 0.$$

Since $\mathrm{D}f(\mathbf{0})$ is a linear map, we have that

$$[\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y) = [\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_x) + [\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_y) = 1.$$

However, by Corollary 2.1.5, we can obtain an expression for $[\mathrm{D}f(\mathbf{0})(\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y)$ by evaluating the directional derivative in the direction of $(1, 1)$.

$$[\mathrm{D}f(\mathbf{0})](\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y) = \sqrt{2}[\mathrm{D}f(\mathbf{0})]\left(\frac{\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y}{\sqrt{2}}\right) = \sqrt{2}\cos^3\theta\bigg|_{\theta = \frac{\pi}{4}} = \sqrt{2}\left(\frac{1}{\sqrt{2}}\right)^3 = \frac{1}{2}.$$

Putting this all together leads to the statement that $\frac{1}{2} = 1$, which gives us our desired contradiction.

An important consideration that should be on the forefront of our minds is whether the way we've defined the notion of derivatives is consistent with the one-variable definition of the derivative. That is, since $\mathbb{R}$ is (one-dimensional) Euclidean space, we can talk about the notion of a map $f : \mathbb{R} \to \mathbb{R}$ being differentiable with respect to Definition 2.1.1. However, at the surface level, this appears to be a different notion to that of f being differentiable in the ordinary, one-variable sense. We must check that these two notions actually coincide, so that our definition of the derivative is indeed a generalisation of the notion of differentiability to the multivariable case.

Proposition 2.1.6. *Let $A \subseteq \mathbb{R}$ be open and let $x_0 \in A$. Then $f : A \to \mathbb{R}$ is differentiable at x_0 if and only if f is differentiable in the one-variable sense, i.e. the following limit exists*

$$f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Moreover, $[Df(x_0)](1) = f'(x_0)$.

Proof. ($\implies$) Suppose f is differentiable at x_0. By Corollary 2.1.5, the directional derivative of f at x_0 in the direction of $1 \in \mathbb{R}$ exists, and is given by

$$[\mathrm{D}f(x_0)](1) = \lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t}.$$

Let $x = x_0 + t$, which will be an element of the domain A provided that t is chosen sufficiently small in magnitude. Then $x \to x_0$ as $t \to 0$, and so we have that

$$[\mathrm{D}f(x_0)](1) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

($\impliedby$) Suppose f is differentiable at x_0 in the one-variable sense. Let $x \in A$. Define the function $R : A \to \mathbb{R}$ by

$$R(x) = f(x) - f(x_0) - f'(x_0)(x - x_0).$$

Then

$$\frac{|R(x)|}{|x - x_0|} = \left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right|.$$

Taking the limit as $x \to x_0$ gives

$$\lim_{x \to x_0} \frac{|R(x)|}{|x - x_0|} = \left| \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| = |f'(x_0) - f'(x_0)| = 0.$$

From the definition of R, we obtain the desired expression

$$f(x) = f(x_0) + [\mathrm{D}f(x_0)](x - x_0) + R(x)$$

where we have defined the linear map $\mathrm{D}f(x_0) : \mathbb{R} \to \mathbb{R}$ by

$$[\mathrm{D}f(x_0)](a) = f'(x_0)a.$$

This shows that f is differentiable at x_0 with derivative $\mathrm{D}f(x_0)$. Evaluating the derivative at $a = 1$ gives $[\mathrm{D}f(x_0)](1) = f'(x_0)$, as required. $\qquad\square$

Given the linear nature of the derivative map, one can recover all of the information about the derivative by evaluating it on a basis of E. This means that there is a distinguished set of directional derivatives that serve to describe the total derivative of the function. These are none other than the function's *partial derivatives*, which are simply the directional derivatives of the function in the direction of the basis vectors.

Definition 2.1.7. Let $\{e_1, \ldots, e_n\}$ be a basis for E. The **partial derivative** of $f : U \to F$ with respect to the ith coordinate at $x_0 \in U$ is the directional derivative of f at x_0 in the direction of the basis vector e_i, i.e.

$$\frac{\partial f}{\partial x_i}(x_0) = \lim_{t \to 0} \frac{f(x_0 + te_i) - f(x_0)}{t}.$$

The definition of a partial derivative closely resembles that of a one-variable derivative. Indeed, we are simply restricting our attention to a one-dimensional subset of the domain (a straight line segment in the direction of the basis vector e_i passing through the point of interest x_0), and so the corresponding limit is a one-dimensional limit. This means that partial derivatives can be readily computed using techniques for computing derivatives of single-variable, real-valued functions. To compute the partial derivative with respect to some coordinate x_i, you essentially treat all other coordinates/variables as constants and consider the function as a one-variable function in x_i, and thus compute the one-variable derivative of the function with respect to x_i.

For example, consider the scalar function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x, y) = 3x^2 y + x \cos(y).$$

Again, we are working in the standard basis, so our basis vectors are

$$\hat{\mathbf{e}}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \hat{\mathbf{e}}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Let's first compute the partial derivative with respect to the x-coordinate directly from the definition. As is conventional, the coordinate x corresponds to the basis vector $\hat{\mathbf{e}}_1$, so the corresponding partial derivative at $\mathbf{x_0} = (x_0, y_0)^T \in \mathbb{R}^2$ is

$$
\begin{aligned}
\frac{\partial f}{\partial x}(\mathbf{x_0}) &= \lim_{t \to 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} \\
&= \lim_{t \to 0} \frac{3(x_0 + t)^2 y + (x_0 + t)\cos(y_0) - 3x_0^2 y_0 - x_0 \cos(y_0)}{t} \\
&= \lim_{t \to 0} \frac{3x_0^2 y + 6x_0 y_0 t + 3t^2 y_0 + x_0 \cos(y_0) + t\cos(y_0) - 3x_0^2 y_0 - x_0 \cos(y_0)}{t} \\
&= \lim_{t \to 0} \frac{6x_0 y_0 t + 3t^2 y_0 + t\cos(y_0)}{t} \\
&= \lim_{t \to 0} \left(6x_0 y_0 + 3t y_0 + \cos(y_0) \right) \\
&= 6x_0 y_0 + \cos(y_0).
\end{aligned}
$$

Observe that we have basically computed a derivative by first principles. The faster approach that makes use of our knowledge of one-variable derivatives goes as follows: fix $y_0 \in \mathbb{R}$ and consider the restricted (now single-variable) function $g : \mathbb{R} \to \mathbb{R}$ given by

$$g(x) = f(x, y_0) = 3x^2 y_0 + x\cos(y_0).$$

We can readily compute the one-variable derivative

$$g'(x) = 6xy_0 + \cos(y_0).$$

But notice that

$$\frac{\partial f}{\partial x}(\mathbf{x_0}) = g'(x_0).$$

We will use the latter method to compute the partial derivative of f with respect to the coordinate y, which corresponds to the basis vector $\hat{\mathbf{e}}_2$. Fix $x_0 \in \mathbb{R}$, and consider the restricted map $h : \mathbb{R} \to \mathbb{R}$ given by

$$h(y) = f(x_0, y) = 3x_0^2 y + x\cos(y).$$

Then the partial derivative with respect to y is given by

$$\frac{\partial f}{\partial y}(\mathbf{x_0}) = h'(y_0) = 3x_0^2 - x_0\sin(y_0).$$

Wasn't that so much easier to do? This is the method we will adopt to compute partial derivatives. After all, why would we want to compute derivatives via first principles given what we know about single-variable calculus? (Answer: we don't!)

Now, as we eluded to previously, the reason why we're making such a big deal about the partial derivatives is because we can express the (total) derivative of a function entirely in terms of its partial derivatives. Since the partial derivatives are fairly easy to compute, this gives us a practical means to compute the total derivative. The foundation of this method will be built upon the following result.

Corollary 2.1.8. *Suppose $f : U \to F$ is differentiable at $x_0 \in U$. Then all of the partial derivatives of f at x_0 exist, and moreover we have*

$$[Df(x_0)](e_i) = \frac{\partial f}{\partial x_i}(x_0)$$

for each $i = 1, \ldots, n$.

Proof. Since, by definition, the partial derivatives of f at x_0 are simply directional derivatives of f at x_0 in the direction of the basis vectors, all of the partial derivatives at x_0 exist by Corollary 2.1.5. Moreover for each $i \in \{1, \ldots, n\}$, we have that

$$[\mathrm{D}f(x_0)](e_i) = \lim_{t\to 0} \frac{f(x_0 + te_i) - f(x_0)}{t} = \frac{\partial f}{\partial x_i}(x_0).$$

$\square$

We are now ready to discuss our method for computing the derivative of a multivariable function f from its partial derivatives. We'll start by considering the case where f is a scalar function then work our way up to a vector-valued function. Suppose we somehow knew that the scalar function $f : U \to \mathbb{R}$ is differentiable at $x_0 \in U$ and want to compute its derivative $\mathrm{D}f(x_0)$. Since $\mathrm{D}f(x_0) \in L(E, \mathbb{R}) = E^*$, i.e. it is a linear functional, we can expand $\mathrm{D}f(x_0)$ in terms of the *dual basis* introduced in Section 1.4

$$\mathrm{D}f(x_0) = \sum_{i=1}^{n} a_i v_i^*$$

for some coefficients $a_i \in \mathbb{R}$. To determine the values of a_i, we evaluate the operators on both sides of the above expression at the kth basis vector e_k

$$\begin{aligned}
[\mathrm{D}f(x_0)](e_k) &= \sum_{i=1}^{n} a_i v_i^*(e_k) \\
&= \sum_{i=1}^{n} a_i \delta_{ik} \\
&= a_k
\end{aligned}$$

but the left-hand side is simply the kth partial derivative of f at x_0, so that

$$a_k = \frac{\partial f}{\partial x_k}(x_0).$$

Hence, the derivative of a scalar function has the following representation in terms of its partial derivatives

$$\mathrm{D}f(x_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)\, v_i^*.$$

Thus, the task of computing the derivative of a differentiable function reduces to computing each of its partial derivatives. The Jacobian matrix of f at x_0, i.e. the matrix representation of $\mathrm{D}f(x_0)$ is simply a $1 \times n$ matrix of the form

$$\begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) & \frac{\partial f}{\partial x_2}(x_0) & \frac{\partial f}{\partial x_3}(x_0) & \cdots & \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix}.$$

Now let's consider the case for vector-valued functions. Let $\{\bar{e}_1, \ldots, \bar{e}_m\}$ be a basis for F, so that for all $x \in U$, we can write

$$f(x) = \sum_{i=1}^{m} f_i(x)\bar{e}_i$$

where each $f_i : U \to \mathbb{R}$ are the component functions of f. Each of the component functions are scalar functions. The differentiability of f implies the differentiability of each of the f_i. In fact, we also have the converse.

Proposition 2.1.9. *A function $f : U \to F$ is differentiable at $x_0 \in U$ if and only if all of its component functions $f_i : U \to \mathbb{R}$ are differentiable at x_0.*

Proof. ($\Longrightarrow$) Suppose f is differentiable at $x_0 \in U$. Then

$$f(x) = f(x_0) + [\mathrm{D}f(x_0)](x - x_0) + R(x)$$

for some $R : U \to F$ satisfying

$$\lim_{\|x - x_0\|} \frac{\|R(x)\|_\infty}{\|x - x_0\|} = 0$$

for all $x \in U$. Notice that we have equipped F with the maximum norm $\|-\|_\infty$, but the final result will be independent of the choice of norm by way of Proposition 2.1.2.

We expand the first expression in terms of the basis in F (i.e. in terms of component functions)

$$\sum_{i=1}^{m} f_i(x)\bar{e}_i = \sum_{i=1}^{m} \Big(f_i(x_0) + [\mathrm{D}f_i(x_0)](x - x_0) + R_i(x) \Big)\bar{e}_i$$

where $\mathrm{D}f_i(x_0)$ and R_i are the ith component functions of $\mathrm{D}f(x_0)$ and R, respectively. Then by linear independence of the $\bar{e}_i$, we can equate the above expression component-wise to obtain

$$f_i(x) = f_i(x_0) + [\mathrm{D}f_i(x_0)](x - x_0) + R_i(x).$$

Recall from Section 1.10 that an arbitrary linear map in $L(E, F)$ can be expressed in terms of a basis consisting of linear maps $\phi_{ij} : E \to F$ satisfying

$$\phi_{ij}(e_k) = \delta_{jk}\bar{e}_i.$$

The corresponding representation of $\mathrm{D}f(x_0)$ in terms of this basis is

$$\mathrm{D}f(x_0) = \sum_{i=1}^{m}\sum_{j=1}^{n} J_{ij}(x_0)\phi_{ij}$$

where $J_{ij}(x_0) \in \mathbb{R}$ is the i,j-element of the Jacobian matrix of f at x_0. Let $x = a_1 e_1 + \cdots + a_n e_n \in E$. Then

$$[\mathrm{D}f(x_0)](x) = \sum_{i=1}^{m} \sum_{j=1}^{n} J_{ij}(x_0) \sum_{k=1}^{n} a_k \phi_{ij}(e_k)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} J_{ij}(x_0) \sum_{k=1}^{n} a_k \delta_{jk} \bar{e}_i$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} J_{ij}(x_0) a_j \right) \bar{e}_i$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} J_{ij}(x_0) v_j^*(x) \right) \bar{e}_i$$

Based on the above representation of $[\mathrm{D}f(x_0)](x)$ in terms of the basis in F, we see that

$$[\mathrm{D}f_i(x_0)](x) = \sum_{j=1}^{n} J_{ij}(x_0) v_j^*(x)$$

which holds for all $x \in E$. The ith component of the derivative $\mathrm{D}f_i(x_0) : E \to \mathbb{R}$ is hence a linear map, since it is a linear combination of linear maps.

Furthermore, for each $i \in \{1, \ldots, m\}$, we have that

$$0 \leq \frac{|R_i(x)|}{\|x - x_0\|_\infty} \leq \max_{k \in \{1, \ldots, m\}} \frac{|R_k(x)|}{\|x - x_0\|_\infty} = \frac{\|R(x)\|_\infty}{\|x - x_0\|_\infty}.$$

Applying the limit as $\|x - x_0\|_\infty \to 0$ and applying the Squeeze Theorem gives

$$\lim_{\|x - x_0\|_\infty \to 0} \frac{|R_i(x)|}{\|x - x_0\|_\infty} = 0.$$

This shows that f_i is differentiable at x_0 with derivative $[\mathrm{D}f(x_0)]_i$ and remainder function R_i.

($\Longleftarrow$) Suppose each of the component functions f_i are differentiable at x_0. Then

$$f_i(x) = f_i(x_0) + [\mathrm{D}f_i(x_0)](x - x_0) + R_i(x)$$

where $R_i : U \to \mathbb{R}$ satisfies

$$\lim_{\|x - x_0\|_\infty \to 0} \frac{|R_i(x)|}{\|x - x_0\|_\infty} = 0$$

for each $i \in \{1, \ldots, m\}$. Define the functions $\mathrm{D}f(x_0) : E \to F$ and $R : U \to F$ by

$$[\mathrm{D}f(x_0)](x) = \sum_{i=1}^{m} [\mathrm{D}f_i(x_0)](x) \bar{e}_i$$

$$R(x) = \sum_{i=1}^{m} R_i(x) \bar{e}_i.$$

Thus, we clearly have that

$$f(x) = f(x_0) + [\mathrm{D}f(x_0)](x - x_0) + R(x).$$

Now, $\mathrm{D}f(x_0)$ is linear because it is a sum of linear maps. We also have that

$$0 \leq \frac{\|R(x)\|_\infty}{\|x - x_0\|_\infty} \leq \sum_{i=1}^{m} \frac{|R_i(x)|}{\|x - x_0\|_\infty}$$

for all $x \in U$. Applying the limit as $\|x - x_0\|_\infty \to 0$ and using the Squeeze Theorem gives

$$\lim_{\|x-x_0\|} \frac{\|R(x)\|_\infty}{\|x - x_0\|} = 0$$

which proves that f is differentiable at x_0 with derivative $\mathrm{D}f(x_0)$ and remainder function R. $\qquad\square$

In light of the above result, given a vector-valued function $f : U \to F$ that is differentiable at $x_0 \in U$, each of the component functions f_i are scalar functions that are differentiable at x_0, and thus they can each be expressed in terms of their partial derivatives

$$\mathrm{D}f_i(x_0) = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x_0)v_j^*.$$

In the proof for Proposition 2.1.9, we saw that the derivative of f can be constructed from the derivatives of the components of f as follows

$$\mathrm{D}f(x_0) = \sum_{i=1}^m \overline{e}_i \mathrm{D}f_i(x_0) = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x_0)\overline{e}_i v_j^* = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x_0)\phi_{ij}.$$

This implies that the derivative $\mathrm{D}f$ is constructed from all of the partial derivatives of the component functions. Written in this way, we see that the i, j-component of the corresponding Jacobian matrix $J_{ij}(x_0)$ is simply

$$J_{ij}(x_0) = \frac{\partial f_i}{\partial x_j}(x_0).$$

Thus, the derivative of a differentiable multivariable function can be completely expressed in terms of its partial derivatives. To put this into practice, recall our early example of the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f(x, y) = \begin{pmatrix} x^2 + y^2 \\ xy \end{pmatrix}.$$

Let's compute its derivative using our described method. First, we'll explicitly write out the component functions of f

$$f_1(x, y) = x^2 + y^2$$
$$f_2(x, y) = xy.$$

Next, we compute all of the partial derivatives at some arbitrary $\mathbf{x_0} = (x_0, y_0)^T \in \mathbb{R}^2$

$$\frac{\partial f_1}{\partial x}(\mathbf{x_0}) = 2x_0 \qquad\qquad \frac{\partial f_1}{\partial y}(\mathbf{x_0}) = 2y_0$$
$$\frac{\partial f_2}{\partial x}(\mathbf{x_0}) = y_0 \qquad\qquad \frac{\partial f_2}{\partial y}(\mathbf{x_0}) = x_0.$$

Finally, we represent the derivative $\mathrm{D}f(\mathbf{x_0}) : \mathbb{R}^2 \to \mathbb{R}^2$ by its Jacobian matrix

$$\mathrm{D}f(\mathbf{x_0}) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(\mathbf{x_0}) & \frac{\partial f_1}{\partial y}(\mathbf{x_0}) \\ \frac{\partial f_2}{\partial x}(\mathbf{x_0}) & \frac{\partial f_2}{\partial y}(\mathbf{x_0}) \end{pmatrix} = \begin{pmatrix} 2x_0 & 2y_0 \\ y_0 & x_0 \end{pmatrix}.$$

You will notice that this coincides with the derivative for f that we plucked out of thin air when we first considered this example.

We remark that while this is an excellent method for actually computing the derivative, it relies on the assumption that f is actually differentiable. Just like with directional derivatives, the partial derivatives of a multivariable function can exist without the function itself actually being differentiable. The existence of the partial derivatives is a necessary condition for differentiability, but not sufficient. At the very least, with what we know now, we can check the differentiability of a function by computing a candidate for the derivative using the partial derivative method, and then directly checking against the definition for differentiability. This is still somewhat awkward, and indeed we will later resolve this inconvenience by using a nice result to establish a criterion for a function to be differentiable based on the behaviour of its partial derivatives. This will be carried out in Section 2.4.

2.2   Properties of the Derivative

In this section, we'll begin to tackle some easy facts about the derivative. All of these are actually generalised versions of their counterparts in one-variable calculus. Namely, the uniqueness of the derivative, differentiability implying continuity, and how differentiation behaves on sums of vector-valued functions and on products of scalar functions. We'll delay the discussion of differentiation of compositions of functions and the Chain Rule for Section 2.3, as it deserves its very own section.

First, we prove the easy fact that if the derivative of a function exists, it is unique. The difference from the one-variable case is that now the derivative (at a point) is a linear map instead of a single number. To check that it is indeed unique, we suppose the existence of another linear map defined on the same domain that satisfy the properties of the derivative of the given function, and show that the two linear maps agree on their entire (shared) domain.

Proposition 2.2.1. *If $f : U \to F$ is differentiable at $x_0 \in U$, the derivative $Df(x_0)$ is unique.*

Proof. Suppose the linear maps $\mathrm{D}f(x_0) \in L(E, F)$ and $T \in L(E, F)$ are both derivatives of f at x_0. Let $y \in E$. Then by Proposition 2.1.3, we have that

$$[\mathrm{D}f(x_0)](y) = \lim_{t \to 0} \frac{f(x_0 + ty) - f(x_0)}{t} = Ty.$$

Since the above equality holds for any $y \in E$, it follows that $\mathrm{D}f(x_0) = T$. $\qquad\qquad \square$

Next, we prove the famous fact that the differentiability of a mapping implies that the mapping is also continuous. So, differentiability is still a stronger property than continuity, even in the multivariable case.

Theorem 2.2.2. *(Differentiability implies continuity) Suppose $f : U \to F$ is differentiable at $x_0 \in U$. Then f is continuous at x_0.*

Proof. Since f is differentiable at x_0, then for all $x \in U$,

$$f(x) - f(x_0) = [\mathrm{D}f(x_0)](x - x_0) + R(x).$$

We can thus establish the following inequality

$$0 \le \|f(x) - f(x_0)\| \le \|[\mathrm{D}f(x_0)](x - x_0)\| + \|R(x)\|.$$

Consider taking the limit as $x \to x_0$ of the above

$$\lim_{x \to x_0} \|f(x) - f(x_0)\| \le \lim_{x \to x_0} \|[\mathrm{D}f(x_0)](x - x_0)\| + \lim_{x \to x_0} \|R(x)\|.$$

Any vector norm $\|-\|$ is a continuous function. Moreover, since $\mathrm{D}f(x_0)$ is a linear map, it is continuous by Theorem 1.7.6. We can hence 'pull the limit inside the function' on the first term on the right-hand-side of the above as follows

$$\lim_{x \to x_0} \|[\mathrm{D}f(x_0)](x - x_0)\| = \left\| [\mathrm{D}f(x_0)]\left(\lim_{x \to x_0} (x - x_0) \right) \right\| = \|[\mathrm{D}f(x_0)](x_0 - x_0)\| = \|[\mathrm{D}f(x_0)](0)\| = 0.$$

This yields the following simplification

$$
\begin{aligned}
0 \le \lim_{x \to x_0} \|f(x) - f(x_0)\| &\le \lim_{x \to x_0} \|R(x)\| \\
&= \lim_{x \to x_0} \left(\|x - x_0\| \frac{\|R(x)\|}{\|x - x_0\|} \right) \\
&= \left(\lim_{x \to x_0} \|x - x_0\| \right) \left(\lim_{x \to x_0} \frac{\|R(x)\|}{\|x - x_0\|} \right) \\
&= 0.
\end{aligned}
$$

It follows that $\|f(x) - f(x_0)\| \to 0$ as $x \to x_0$. This implies that f is continuous at x_0. $\qquad \square$

Given a differentiable function $f : U \to F$, is the derivative map $\mathrm{D}f : U \to L(E, F)$ continuous? The answer is not necessarily. In fact, we don't even need to invoke a multivariable function to construct a suitable counterexample. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Since f is differentiable on $\mathbb{R}$ in the one-variable sense, it is differentiable in the multivariable sense. Its corresponding differentiation map $\mathrm{D}f : \mathbb{R} \to L(\mathbb{R}, \mathbb{R})$ is

$$\mathrm{D}f(x) = f'(x)\, I_1 = \begin{cases} \left(2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)\right) I_1 & x \neq 0 \\ 0\, I_1 & x = 0. \end{cases}$$

We claim that $\mathrm{D}f$ is *not* continuous at $x = 0$. We will show this by showing that $\mathrm{D}f$ is not sequentially continuous at $x = 0$. That is, we will identify a real-valued sequence $(x_n)_{n=1}^\infty$ that converges to 0, but orchestrate for $\|\mathrm{D}f(x_n) - \mathrm{D}f(0)\|$ to not converge to 0 (in operator norm). To this end, consider the sequence defined by

$$x_n = \frac{1}{2\pi n}.$$

Clearly, we have that $x_n \to 0$ as $n \to \infty$. Moreover, we have for all $n \in \mathbb{N}$

$$\|\mathrm{D}f(x_n) - \mathrm{D}f(0)\| = \|(-\cos(2\pi n))I_1\| = 1.$$

Thus, the derivative map $\mathrm{D}f$ of a differentiable function is not necessarily continuous. Those that do indeed have a continuous derivative fall into a class of functions called $\mathcal{C}^1$ functions. Such functions will be the topic of Section 2.4.

Now, we will consider ways in which we may obtain new differentiable functions from old. One of the simplest ways to do so is to consider a vector sum of two differentiable functions. The theorem below says that the resulting sum function is itself differentiable, and the formula for the derivative is simply the sum of the respective derivatives.

Theorem 2.2.3. *(Sum Rule) Suppose the maps $f : U \to F$ and $g : U \to F$ are differentiable at $x_0 \in U$. Then the sum function $f + g : U \to F$ defined by*

$$(f + g)(x) = f(x) + g(x)$$

is differentiable at x_0, and its derivative is given by

$$D(f + g)(x_0) = Df(x_0) + Dg(x_0).$$

Proof. From the differentiability of f and g at x_0, we can establish the existence of the mappings $\mathrm{D}f(x_0), \mathrm{D}g(x_0) \in L(E, F)$ and $R_f, R_g : U \to F$ satisfying

$$f(x) = f(x_0) + [\mathrm{D}f(x_0)](x - x_0) + R_f(x)$$
$$g(x) = g(x_0) + [\mathrm{D}g(x_0)](x - x_0) + R_g(x)$$

and

$$\lim_{\|x - x_0\| \to 0} \frac{\|R_f(x)\|}{\|x - x_0\|} = 0$$
$$\lim_{\|x - x_0\| \to 0} \frac{\|R_g(x)\|}{\|x - x_0\|} = 0.$$

Summing the expressions for $f(x)$ and $g(x)$ yields

$$(f + g)(x) = (f + g)(x_0) + (\mathrm{D}f(x_0) + \mathrm{D}g(x_0))(x - x_0) + (R_f(x) + R_g(x)).$$

Note that $(\mathrm{D}f(x_0) + \mathrm{D}g(x_0)) \in L(E, F)$, since the sum of linear maps is linear. Moreover, if we define $R_{f+g} : U \to F$ by $R_{f+g} = R_f + R_g$, then

$$0 \leq \lim_{\|x-x_0\| \to 0} \frac{\|R_{f+g}(x)\|}{\|x - x_0\|} \leq \lim_{\|x-x_0\| \to 0} \left(\frac{\|R_f(x)\|}{\|x - x_0\|} + \frac{\|R_g(x)\|}{\|x - x_0\|} \right) = 0.$$

This proves that $f + g$ is differentiable at x_0. Moreover, we see that $\mathrm{D}(f+g)(x_0) = \mathrm{D}f(x_0) + \mathrm{D}g(x_0)$. $\square$

Next, we consider a multivariable product rule for differentiation. Given two differentiable scalar-valued functions (i.e. the outputs are real numbers), we can sensibly form the product function and think about trying to differentiate it. The Product Rule below says that the differentiation procedure on the product is A-OK and furthermore there is a formula for the derivative that is a mere generalisation of the familiar one-variable product rule.

Theorem 2.2.4. *(Product Rule) Suppose the scalar functions $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}$ are differentiable at $x_0 \in U$. Then the product function $fg : U \to \mathbb{R}$ defined by*

$$(fg)(x) = f(x)g(x)$$

is differentiable at x_0, and its derivative is given by

$$D(fg)(x_0) = f(x_0)Dg(x_0) + g(x_0)Df(x_0).$$

Proof. From the differentiability of f and g at x_0, we can establish the existence of the mappings $\mathrm{D}f(x_0), \mathrm{D}g(x_0) \in L(E, F)$ and $R_f, R_g : U \to F$ satisfying

$$f(x) = f(x_0) + [\mathrm{D}f(x_0)](x - x_0) + R_f(x)$$
$$g(x) = g(x_0) + [\mathrm{D}g(x_0)](x - x_0) + R_g(x)$$

and

$$\lim_{\|x-x_0\| \to 0} \frac{\|R_f(x)\|}{\|x - x_0\|} = 0$$
$$\lim_{\|x-x_0\| \to 0} \frac{\|R_g(x)\|}{\|x - x_0\|} = 0.$$

Multiplying the expressions for $f(x)$ and $g(x)$ yields

$$(fg)(x) = (fg)(x_0) + f(x_0)[\mathrm{D}g(x_0)](x - x_0) + f(x_0)R_g(x) + g(x_0)[\mathrm{D}f(x_0)](x - x_0) + \dots$$
$$+ ([\mathrm{D}f(x_0)](x - x_0))([\mathrm{D}g(x_0)](x - x_0)) + R_g(x)[\mathrm{D}f(x_0)](x - x_0) + \dots$$
$$+ g(x_0)R_f(x) + R_f(x)[\mathrm{D}g(x_0)](x - x_0) + R_f(x)R_g(x).$$

We will define the linear map $\mathrm{D}(fg)(x_0) \in L(E, \mathbb{R})$ by the linear combination of linear maps

$$\mathrm{D}(fg)(x_0) = f(x_0)\mathrm{D}g(x_0) + g(x_0)\mathrm{D}f(x_0)$$

and define the remainder function $R_{fg} : U \to \mathbb{R}$ by 'all of the other stuff in the expression'

$$R_{fg}(x) = f(x_0)R_g(x) + ([\mathrm{D}f(x_0)](x - x_0))([\mathrm{D}g(x_0)](x - x_0)) + R_g(x)[\mathrm{D}f(x_0)](x - x_0) + \dots$$
$$+ g(x_0)R_f(x) + R_f(x)[\mathrm{D}g(x_0)](x - x_0) + R_f(x)R_g(x).$$

Thus, $(fg)(x)$ can be expressed as follows

$$(fg)(x) = (fg)(x_0) + \mathrm{D}(fg)(x_0)(x - x_0) + R_{fg}(x).$$

It remains to be checked that

$$\lim_{\|x-x_0\| \to 0} \frac{\|R_{fg}(x)\|}{\|x - x_0\|} = 0.$$

By inspection of the expression for R_{fg}, by taking the norm and dividing through by $\|x - x_0\|$, we can repeatedly apply the triangle inequality to separate the normed expression into the norms of each of its individual summands. Any terms containing R_f or R_g can thus be made to vanish by taking the limit $\|x - x_0\| \to 0$ by hypothesis. The only other term to worry about is

$$\frac{\|([\mathrm{D}f(x_0)](x - x_0))([\mathrm{D}g(x_0)](x - x_0))\|}{\|x - x_0\|} \leq \frac{\|\mathrm{D}f(x_0)\|\|x - x_0\|\|\mathrm{D}g(x_0)\|\|x - x_0\|}{\|x - x_0\|}$$
$$= \|\mathrm{D}f(x_0)\|\|\mathrm{D}g(x_0)\|\|x - x_0\|$$

where $\|\mathrm{D}f(x_0)\|$ and $\|\mathrm{D}g(x_0)\|$ are the operator norms of the linear maps $\mathrm{D}f(x_0)$ and $\mathrm{D}g(x_0)$, respectively. Taking the limit as $\|x - x_0\| \to 0$ makes the right-hand side of the above expression vanish, which implies that the left-hand side also vanishes by the Squeeze Theorem. This proves the desired result, and thus demonstrates the differentiability of fg at x_0 with the anticipated derivative. $\qquad \square$

We saw earlier that first-order monomials (e.g. functions of the form $f(x, y, z) = x$, $g(x, y, z) = y$, $h(x, y, z) = z$) are differentiable. Thus, by the Sum Rule and Product Rule together, we have that every multivariable scalar function that is a polynomial in terms of its coordinate variables is differentiable.

With the previous results under our belt, you may be eager to extend the familiar Quotient Rule to multivariable calculus. We will actually delay the proof of it till the following section, where it will be shown to be a straightforward consequence of the Product Rule and Chain Rule.

2.3 Compositions of Differentiable Mappings

Theorem 2.3.1. *(Chain Rule) Let E, F, G be Euclidean spaces, let $U \subseteq E$ and $V \subseteq F$ be open and consider the mappings*

$$f : U \to F$$
$$g : V \to G.$$

such that $f(U) \subseteq V$. If f is differentiable at $x_0 \in U$ and g is differentiable at $f(x_0) \in V$, then the composite mapping

$$g \circ f : U \to G$$

is differentiable at x_0 with derivative $Dg(f(x_0)) \circ Df(x_0)$.

Proof. From the differentiability of f at x_0, we can establish the existence of the mappings $\mathrm{D}f(x_0) \in L(E, F)$ and $R_f : U \to F$ satisfying

$$f(x) = f(x_0) + [\mathrm{D}f(x_0)](x - x_0) + R_f(x)$$

for all $x \in E$, and

$$\lim_{\|x - x_0\| \to 0} \frac{\|R_f(x)\|}{\|x - x_0\|} = 0.$$

Similarly, from the differentiability of g at $f(x_0)$, we can establish the existence of the mappings $\mathrm{D}g(f(x_0)) \in L(F, G)$ and $R_g : V \to G$ satisfying

$$g(y) = g(f(x_0)) + [\mathrm{D}g(f(x_0))](y - f(x_0)) + R_g(y)$$

for all $y \in F$, and

$$\lim_{\|y - f(x_0)\| \to 0} \frac{\|R_g(y)\|}{\|y - f(x_0)\|} = 0.$$

Let $x \in E$. Then $f(x) \in V$ by hypothesis so that $f(x) \in \mathcal{D}(g)$. Thus we can meaningfully substitute $y = f(x)$ into the expression for $g(y)$ to obtain

$$g(f(x)) = g(f(x_0)) + [\mathrm{D}g(f(x_0))](f(x) - f(x_0)) + R_g(f(x)).$$

Now substitute the expression for $f(x)$ into the above

$$g \circ f(x) = g \circ f(x_0) + [\mathrm{D}g(f(x_0))]([\mathrm{D}f(x_0)](x - x_0) + R_f(x)) + R_g(f(x))$$
$$= g \circ f(x_0) + [\mathrm{D}g(f(x_0))] \circ [\mathrm{D}f(x_0)](x - x_0) + R_{g \circ f}(x)$$

where $R_{g \circ f} : U \to G$ is defined as

$$R_{g \circ f}(x) = [\mathrm{D}g(f(x_0))](R_f(x)) + R_g(f(x)).$$

From the above, we can establish the following inequality

$$\frac{\|R_{g \circ f}(x)\|}{\|x - x_0\|} \leq [\mathrm{D}g(f(x_0))] \left(\frac{\|R_f(x)\|}{\|x - x_0\|} \right) + \frac{\|R_g(f(x))\|}{\|x - x_0\|}.$$

Taking the limit as $\|x - x_0\| \to 0$ yields

$$\lim_{\|x-x_0\| \to 0} \frac{\|R_{g \circ f}(x)\|}{\|x - x_0\|} \leq \lim_{\|x-x_0\| \to 0} [\mathrm{D}g(f(x_0))] \left(\frac{\|R_f(x)\|}{\|x - x_0\|} \right) + \lim_{\|x-x_0\| \to 0} \left(\frac{\|R_g(f(x))\|}{\|f(x) - f(x_0)\|} \times \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \right)$$

$$= [\mathrm{D}g(f(x_0))] \left(\lim_{\|x-x_0\| \to 0} \frac{\|R_f(x)\|}{\|x - x_0\|} \right) + \lim_{\|x-x_0\| \to 0} \frac{\|R_g(f(x))\|}{\|f(x) - f(x_0)\|} \frac{\|[\mathrm{D}f(x_0)](x - x_0) + R_f(x)\|}{\|x - x_0\|}$$

$$\leq [\mathrm{D}g(f(x_0))](0) + 0 \times \lim_{\|x-x_0\| \to 0} \left(\|\mathrm{D}f(x_0)\| + \frac{\|R_f(x)\|}{\|x - x_0\|} \right)$$

$$= 0.$$

We remark that the above calculations are valid provided that $f(x) \neq f(x_0)$, so that the expression containing $1/\|f(x) - f(x_0)\|$ is non-singular. However, notice that for any $x \in U$ satisfying $f(x) = f(x_0)$, we have that[1] $R_g(f(x)) = R_g(f(x_0)) = 0$, so the problematic term can be replaced with 0 any time we have $f(x) = f(x_0)$ whilst taking the limit $x \to x_0$.

The above calculation, together with the fact the composition of linear maps is linear, shows that $g \circ f$ is differentiable at x_0 with the expected derivative. $\qquad\square$

Multiplication of Jacobian matrices.

Theorem 2.3.2. *(Quotient Rule) Suppose the scalar functions $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}$ are differentiable at $x_0 \in U$, and moreover we have that $g(x_0) \neq 0$. Then the quotient function $f/g : U \to \mathbb{R}$ defined by*

$$(f/g)(x) = \frac{f(x)}{g(x)}$$

is differentiable at x_0, and its derivative is given by

$$\mathrm{D}(f/g)(x_0) = \frac{g(x_0)\mathrm{D}f(x_0) - f(x_0)\mathrm{D}g(x_0)}{(g(x_0))^2}.$$

Proof. Define the function $h : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ by

$$h(x) = \frac{1}{x}.$$

Then $1/g = h \circ g$. Since h is differentiable on $\mathbb{R}\backslash\{0\}$, and g is differentiable at x_0, it follows that $h \circ g$ is differentiable at x_0 by the Chain Rule, which yields the formula

$$\mathrm{D}(h \circ g)(x_0) = \mathrm{D}h(g(x_0)) \circ \mathrm{D}g(x_0).$$

---

[1]To see this, consider substituting $y = f(x_0)$ into

$$g(y) = g(f(x_0)) + [\mathrm{D}g(f(x_0))](y - f(x_0)) + R_g(y)$$

and solve for $R_g(f(x_0))$.

We recall from one-variable calculus the derivative of $1/x$

$$\mathrm{D}h(g(x_0)) = h'(g(x_0))I_1 = -\frac{1}{(g(x_0))^2}I_1$$

where $I_1 : \mathbb{R} \to \mathbb{R}$ is the 1D identity map $I_1(x) = x$, which has been inserted to make clear that the above expression is an operator expression. Substituting this into our expression for the derivative of $h \circ g$ yields

$$\mathrm{D}(h \circ g)(x_0) = -\frac{1}{(g(x_0))^2}\mathrm{D}g(x_0).$$

Finally, we apply the Product Rule to f and $1/g = h \circ g$ and conclude that f/g is differentiable at x_0 with derivative

$$\begin{aligned}\mathrm{D}(\,f/g\,)(x_0) &= f(x_0)\mathrm{D}(h \circ g)(x_0) + (h \circ g)(x_0)\mathrm{D}f(x_0) \\ &= \frac{-f(x_0)\mathrm{D}g(x_0)}{(g(x_0))^2} + \frac{\mathrm{D}f(x_0)}{g(x_0)} \\ &= \frac{g(x_0)\mathrm{D}f(x_0) - f(x_0)\mathrm{D}g(x_0)}{(g(x_0))^2}.\end{aligned}$$

$\square$

An important application of the Chain Rule is the computation of derivatives of functions with respect to an alternative coordinate system. By default, when we work in $\mathbb{R}^n$, we express every vector in terms of an orthonormal basis $\{\hat{\mathbf{e}}_1, \ldots, \hat{\mathbf{e}}_n\}$ in $\mathbb{R}^n$ (typically the standard basis). So given an arbitrary vector $\mathbf{x} \in \mathbb{R}^n$ and the representation

$$\mathbf{x} = \sum_{i=1}^{n} x_i\hat{\mathbf{e}}_i,$$

the n-tuple $(x_1, \ldots, x_n)$ of real numbers are the coordinates of $\mathbf{x}$ in a *rectangular* or *Cartesian* coordinate system. The name 'rectangular' comes from the fact that it is very easy to specify rectangular regions (i.e. n-dimensional boxes) using this coordinate system; they are simply the set of vectors whose coordinates lie in a Cartesian product of intervals

$$\{\mathbf{x} \in \mathbb{R}^n \mid (x_1, \ldots, x_n) \in [a_1, b_1] \times \cdots \times [a_n, b_n]\}$$

where $a_i, b_i \in \mathbb{R}$.

Sometimes, it is convenient to express vectors in different coordinate systems. One reason is to exploit a particular symmetry that a function may exhibit that is not too clear to realise in Cartesian coordinates. Another reason is that the domain of the function may not be a rectangular region (or the entire vector space), and so it becomes easier to specify the vectors that lie in the domain using different coordinates.

To change from Cartesian coordinates to a different coordinate system, one applies a suitable *coordinate transformation T*. To express a mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ with respect to our new coordinate system, we simply take the composition $f \circ T$. Thus, to compute the derivative of f with respect to the new coordinates, we differentiate $f \circ T$. This is where the Chain Rule comes in.

Let's consider a concrete example. A popular choice of coordinates in $\mathbb{R}^2$ is the *plane polar coordinate system*, in which every $\mathbf{x} \in \mathbb{R}^2$ is specified by its distance r from the origin, and the angle θ it makes with the positive x-axis (i.e. with the basis vector $\hat{\mathbf{e}}_1$), measured in a counterclockwise direction. That is, any vector can be represented in terms of the coordinates (r, θ), where the corresponding representation is given by the coordinate transformation $T : [0, \infty) \times [0, 2\pi) \to \mathbb{R}^2$ given by

$$T(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}.$$

This coordinate transformation is surjective, i.e. every vector in $\mathbb{R}^2$ can be represented in plane polar coordinates.

Now, let's consider the scalar function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x, y) = x^2 + y^2.$$

We can compute the derivative of f with respect to plane polar coordinates using the Chain Rule.

$$
\begin{aligned}
\mathrm{D}(f \circ T)(r, \theta) &= \mathrm{D}f(T(r, \theta)) \circ \mathrm{D}T(r, \theta) \\
&= \left(\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right)\Bigg|_{(x,y)=(r\cos\theta, r\sin\theta)} \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} \\
&= \left(2x \quad 2y \right)\Bigg|_{(x,y)=(r\cos\theta, r\sin\theta)} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \\
&= \left(2r\cos\theta \quad 2r\sin\theta \right) \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \\
&= \left(2r\cos^2\theta + 2r\sin^2\theta \quad -2r^2\sin\theta\cos\theta + 2r^2\sin\theta\cos\theta \right) \\
&= \left(2r \quad 0 \right)
\end{aligned}
$$

Since $\mathrm{D}(f \circ T)(r, \theta) = \left(\frac{\partial(f \circ T)}{\partial r} \quad \frac{\partial(f \circ T)}{\partial \theta} \right)$, by comparison with the above expression we see that

$$\frac{\partial(f \circ T)}{\partial r} = 2r$$

$$\frac{\partial(f \circ T)}{\partial \theta} = 0.$$

Of course, we still have the interpretation that the partial derivatives are the infinitesimal change of the function with respect to a single coordinate, holding all others fixed. So for instance, $\frac{\partial(f \circ T)}{\partial r}$ is the infinitesimal rate of change of f in the radial direction $\hat{\mathbf{r}}$. Which way does the radial direction point? If we perturb r slightly whilst holding the other coordinate, θ, fixed, the x-coordinate changes by an amount $\frac{\partial x}{\partial r}$, and the y-coordinate changes by an amount $\frac{\partial y}{\partial r}$, so a vector $\vec{\mathbf{r}}$ pointing in the radial direction is given by

$$\vec{\mathbf{r}} = \frac{\partial x}{\partial r} \hat{\mathbf{e}}_1 + \frac{\partial y}{\partial r} \hat{\mathbf{e}}_2 = \cos\theta \hat{\mathbf{e}}_1 + \sin\theta \hat{\mathbf{e}}_2.$$

To obtain a unit vector $\hat{\mathbf{r}}$, we then normalise the length of $\vec{\mathbf{r}}$

$$\hat{\mathbf{r}} = \frac{\vec{\mathbf{r}}}{\|\vec{\mathbf{r}}\|} = \cos\theta \hat{\mathbf{e}}_1 + \sin\theta \hat{\mathbf{e}}_2.$$

Here, it so happened that $\vec{\mathbf{r}}$ already had unit length. Notice that the direction of $\hat{\mathbf{r}}$ changes depending on where you are in $\mathbb{R}^2$, as opposed to the direction of the basis vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$, whose directions are the same everywhere. The unit vector $\hat{\mathbf{r}}$ always points radially away from the origin, as expected.

We can play a similar game to compute the unit vector $\hat{\boldsymbol{\theta}}$

$$\hat{\boldsymbol{\theta}} = \left(\frac{\partial x}{\partial \theta} \hat{\mathbf{e}}_1 + \frac{\partial y}{\partial \theta} \hat{\mathbf{e}}_2 \right) \left(\left(\frac{\partial x}{\partial \theta} \right)^2 + \left(\frac{\partial y}{\partial \theta} \right)^2 \right)^{-1/2} = \frac{1}{r} \left(-r\sin\theta \hat{\mathbf{e}}_1 + r\cos\theta \hat{\mathbf{e}}_2 \right) = -\sin\theta \hat{\mathbf{e}}_1 + \cos\theta \hat{\mathbf{e}}_2.$$

For some $\mathbf{x} = (r\cos\theta, r\sin\theta)^T \in \mathbb{R}^2$, $\hat{\boldsymbol{\theta}}$ is tangential to a circle of radius r centred at the origin (pointing in a counterclockwise direction). This aligns with our intuition that θ is an angular coordinate.

Now, notice that the partial derivative of $f \circ T$ with respect to θ vanishes, which reveals that f possesses *angular symmetry*, i.e. the value of f is constant on any circle centred on the origin. Thus, the value of f at

a point $\mathbf{x}$ is only dependent on the distance r of $\mathbf{x}$ from the origin. This is an instance of a property of a function that becomes much easier to notice when expressed in a more appropriate coordinate system.

We can also observe this by expressing f directly in plane polar coordinates

$$f \circ T(r, \theta) = f(r \cos \theta, r \sin \theta) = r^2 \cos^2 \theta + r^2 \sin^2 \theta = r^2.$$

Notice that $f \circ T$ is independent of the angular coordinate θ, which is consistent with our previous observations.

When expressing functions in different coordinate systems, it is common to adopt the abuse of notation by dropping the coordinate transformation T. That is, one would write $f(r, \theta)$ to mean $f \circ T(r, \theta)$. To express f in Cartesian coordinates, one would simply write $f(x, y)$. So, the coordinates appearing as arguments of the function indicate indicate which particular coordinate system we are in. Here, the coordinate pair (r, θ) is implicitly understood to be plane polar coordinates and (x, y) is understood to designate Cartesian coordinates. For the sake of clarity, we will *not* adopt such conventions, but warn the reader that such (mal)practice is standard.

2.4   Mappings of Class $\mathcal{C}^1$

Definition 2.4.1. A mapping $f : U \to F$ is said to be of **class $\mathcal{C}^1$** if it is differentiable on U and its derivative map $Df : U \to L(E, F)$ is continuous on U.

Theorem 2.4.2. *A scalar function $f : U \to \mathbb{R}$ is $\mathcal{C}^1$ if and only if all of the partial derivatives of f with respect to the coordinates of an orthonormal basis in E exist and are continuous on U.*

Proof. ($\implies$) Suppose f is $\mathcal{C}^1$. So f is differentiable and its derivative map $Df : U \to E^*$ is continuous. Recall that Df has the following representation in terms of its partial derivative mappings

$$Df(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) v_i^*$$

where $v_i^* : E \to \mathbb{R}$ is the dual basis vector corresponding to the ith coordinate.

Let $k \in \{1, \ldots, n\}$ and let $x_0 \in U$. Then for all $x \in U$, we have that

$$
\begin{aligned}
0 &\leq \left| \frac{\partial f}{\partial x_k}(x) - \frac{\partial f}{\partial x_k}(x_0) \right| \\
&\leq \left\| \frac{\partial f}{\partial x_k}(x) - \frac{\partial f}{\partial x_k}(x_0) \right\|_\infty \\
&= \| Df(x) - Df(x_0) \|.
\end{aligned}
$$

Thus, by taking the limit as $x \to x_0$ in the above inequality, we see that $\| Df(x) - Df(x_0) \| \to 0$ since Df is continuous at x_0, and thus by the Squeeze Theorem we have that

$$\lim_{x \to x_0} \left| \frac{\partial f}{\partial x_k}(x) - \frac{\partial f}{\partial x_k}(x_0) \right| = 0$$

which shows that $\frac{\partial f}{\partial x_k}$ is continuous at x_0.

($\impliedby$) Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis on E and suppose that all of the partial derivative mappings of f with respect to the coordinates of $\{e_1, \ldots, e_n\}$ exist and are continuous on U.
Let $x_0 = \sum_{i=1}^{n} a_i e_i \in U$.
Let $\epsilon > 0$.

By continuity of the partial derivatives, for each $i \in \{1, \ldots, n\}$, there exists a $\delta_i > 0$ such that for all[2] $x \in B_\infty(x_0, \delta_i)$, we have

$$\left| \frac{\partial f(x)}{\partial x_i} - \frac{\partial f(x_0)}{\partial x_i} \right| < \frac{\epsilon}{n}.$$

Choose $\delta = \min_{i \in \{1, \ldots, n\}} \delta_i$, and let $x = \sum_{i=1}^{n} b_i e_i \in B_\infty(x_0, \delta)$. Now, for each $k \in \{1, \ldots, n-1\}$, we define the point $x_k \in E$ by

$$x_k = \sum_{i=k+1}^{n} a_i e_i + \sum_{i=1}^{k} b_i e_i$$

i.e. the finite sequence $x_0, x_1, x_2, \ldots, x_{n-1}, x_n = x$ gradually replaces the coordinates of x_0 one component at a time with that of x. We have that $x_k \in B_\infty(x_0, \delta)$ for each k, since

$$\|x_k - x_0\|_\infty = \max_{i \in \{1, \ldots, k\}} |b_i - a_i| \leq \max_{i \in \{1, \ldots, n\}} |b_i - a_i| = \|x - x_0\|_\infty < \delta.$$

Observe that we can express $f(x) - f(x_0)$ as the following telescoping sum

$$f(x) - f(x_0) = \sum_{i=1}^{n} f(x_i) - f(x_{i-1}).$$

For each $k \in \{1, \ldots, n\}$, we consider the restriction of f to the line segment connecting x_{k-1} and x_k, which defines a one-variable function $g_k : [0, b_k - a_k] \to \mathbb{R}$ by[3]

$$g_k(t) = f(x_0 + t e_k).$$

We see that g_k is a differentiable one-variable function on $[0, b_k - a_k]$, since the derivative is simply given by the partial derivative of f in the kth coordinate

$$g_k'(t) = \frac{\partial f(x)}{\partial x_k} \bigg|_{x = x_0 + t e_k}.$$

Hence, by the Mean Value Theorem from one-variable calculus, there exists a $c_k \in [0, b_k - a_k]$ such that

$$g_k'(c_k) = \frac{g_k(b_k - a_k) - g_k(0)}{b_k - a_k}$$

i.e. we have that

$$\frac{\partial f(x_0 + c_k e_k)}{\partial x_k}(b_k - a_k) = f(x_k) - f(x_{k-1}).$$

Putting this all together, we have that

$$
\begin{aligned}
f(x) - f(x_0) &= \sum_{i=1}^{n} \frac{\partial f(x_0 + c_i e_i)}{\partial x_i}(b_i - a_i) \\
&= \sum_{i=1}^{n} \frac{\partial f(x_0 + c_i e_i)}{\partial x_i} v_i^*(x - x_0) \\
&= \sum_{i=1}^{n} \left[\frac{\partial f(x_0)}{\partial x_i} - \left(\frac{\partial f(x_0)}{\partial x_i} - \frac{\partial f(x_0 + c_i e_i)}{\partial x_i} \right) \right] v_i^*(x - x_0) \\
&= [\mathrm{D}f(x_0)](x - x_0) + R(x)
\end{aligned}
$$

---

[2]By $B_\infty(x_0, \delta_i)$, we mean the open ball with centre x_0 and radius δ_i constructed from the metric induced by the maximum norm.

[3]Without loss of generality, we assume that $a_k < b_k$. If $b_k < a_k$, we construct g_k on the interval $[b_k - a_k, 0]$. When we apply the Mean Value Theorem, we pick up a minus sign from both the numerator and denominator, which cancel, giving us the same expression as before.

where

$$R(x) = \sum_{i=1}^{n} \left(\frac{\partial f(x_0 + c_i e_i)}{\partial x_i} - \frac{\partial f(x_0)}{\partial x_i} \right)(b_i - a_i).$$

Taking absolute values and applying the triangle inequality, we have that

$$|R(x)| \leq \sum_{i=1}^{n} \left| \frac{\partial f(x_0 + c_i e_i)}{\partial x_i} - \frac{\partial f(x_0)}{\partial x_i} \right| |b_i - a_i|$$

$$\leq \sum_{i=1}^{n} \left| \frac{\partial f(x_0 + c_i e_i)}{\partial x_i} - \frac{\partial f(x_0)}{\partial x_i} \right| \|x - x_0\|_\infty$$

$$< \epsilon \|x - x_0\|_\infty$$

i.e. we have that

$$\frac{|R(x)|}{\|x - x_0\|_\infty} < \epsilon.$$

This shows that f is differentiable at x_0, with derivative

$$\mathrm{D}f(x_0) = \sum_{i=1}^{n} \frac{\partial f(x_0)}{\partial x_i} v_i^*.$$

Since the choice of $x_0 \in U$ was arbitrary, it follows that f is differentiable on U. Furthermore, since all of the partial derivative maps are continuous functions on U, it follows that $\mathrm{D}f$ is continuous because it a sum of continuous maps. Hence, f is $\mathcal{C}^1$. $\qquad\square$

Theorem 2.4.3. *A mapping $f : U \to F$ is $\mathcal{C}^1$ if and only if all of the partial derivatives of f with respect to the coordinates of an orthonormal basis in E exist and are continuous on U.*

Proof. ($\implies$) Suppose f is $\mathcal{C}^1$. Thus, f is differentiable, and its derivative $\mathrm{D}f : U \to L(E, F)$ is continuous. This implies that all of the component functions (with respect to some orthonormal basis) $f_i : U \to \mathbb{R}$ are differentiable scalar functions by Proposition 2.1.9, and the components of the derivative $\mathrm{D}f_i : U \to E^*$ are continuous mappings by Proposition 1.7.5. Thus, each of the component functions f_i are $\mathcal{C}^1$. We then apply Theorem 2.4.2 to conclude that for each $i \in \{1, \ldots, m\}$, the partial derivative mappings

$$\frac{\partial f_i}{\partial x_j} : U \to \mathbb{R}$$

exist and are continuous for each $j = 1, \ldots, n$.

($\impliedby$) Suppose that all of the partial derivative mappings of f with respect to the coordinates of an orthonormal basis in E exist and are continuous on U. We can apply Theorem 2.4.2 to conclude that the component function $f_i : U \to \mathbb{R}$ is $\mathcal{C}^1$ for each $i \in \{1, \ldots, m\}$. So each f_i is differentiable, and each component of the derivative $\mathrm{D}f_i : U \to E^*$ is continuous. Thus, by Proposition 2.1.9, we see that f is differentiable, and from Proposition 1.7.5, we see that $\mathrm{D}f : U \to L(E, F)$ is continuous. This shows that f is $\mathcal{C}^1$. $\qquad\square$

2.5 Higher Order Derivatives

Definition 2.5.1. A mapping $f : U \to F$ is said to be of **class $\mathcal{C}^k$** (for some $k \in \mathbb{N}$) if it is k-times differentiable on U and its derivative mappings $\mathrm{D}^i f : U \to L(E, F)$ are continuous on U for all orders i up to k (i.e. for all $1 \leq i \leq k$).

Theorem 2.5.2. *(Clairaut-Schwarz Theorem) Suppose $f : U \to \mathbb{R}$ is $\mathcal{C}^2$. Then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

for all $i, j \in \{1, \ldots, n\}$.

Definition 2.5.3. A mapping $f : U \to F$ is said to be of **class $\mathcal{C}^\infty$** or **smooth** if it is $\mathcal{C}^k$ for all $k \in \mathbb{N}$.

2.6   Taylor's Theorem

2.7   Optimisation

Definition 2.7.1. Let $f : U \to F$ be differentiable. A point $x_0 \in U$ is called a **critical point** of f if $\mathrm{D}f(x_0) = 0$.

Definition 2.7.2. Let $f : U \to \mathbb{R}$ be a scalar function.

(a) A point $x_0 \in U$ is called a **local minimum** of f if there exists a neighbourhood A of x_0 such that $f(x_0) \leq f(x)$ for all $x \in A$.

(b) A point $x_0 \in U$ is called a **local maximum** of f if there exists a neighbourhood A of x_0 such that $f(x) \leq f(x_0)$ for all $x \in A$.

(c) A point $x_0 \in U$ is called a **local extremum** if it is either a local minimum or local maximum.

Theorem 2.7.3. *Let $f : U \to \mathbb{R}$ be differentiable, and suppose $x_0 \in U$ is a local extremum. Then x_0 is a critical point of f.*

Definition 2.7.4. Let V be a (real) vector space. A mapping $g : V \times V \to \mathbb{R}$ is called a **bilinear form** if for all $x, y, z \in V$ and $a \in \mathbb{R}$, we have

(a) $g(x + y, z) = g(x, z) + g(y, z)$.

(b) $g(ax, y) = ag(x, y)$.

(c) $g(x, y + z) = g(x, y) + g(x, z)$.

(d) $g(x, ay) = ag(x, y)$.

Definition 2.7.5. Let V be a vector space. A bilinear form $g : V \times V \to \mathbb{R}$ is called **symmetric** if $g(x, y) = g(y, x)$ for all $x, y \in V$.

Definition 2.7.6. Let V be a vector space and let $g : V \times V \to \mathbb{R}$ be a symmetric bilinear form. The **quadratic form** associated with g is the mapping $Q : V \to \mathbb{R}$ defined by

$$Q(x) = g(x, x).$$

Definition 2.7.7. Let V be a vector space, and let $Q : V \to \mathbb{R}$ be a quadratic form.

(a) We say that Q is **positive definite** if $Q(x) > 0$ for all $x \in V \setminus \{0\}$.

(b) We say that Q is **negative definite** if $Q(x) < 0$ for all $x \in V \setminus \{0\}$.